



## Research papers

## Spatio-temporal modelling of the status of groundwater droughts

B.P. Marchant<sup>a,\*</sup>, J.P. Bloomfield<sup>b</sup><sup>a</sup> British Geological Survey, Keyworth NG12 5GG, UK<sup>b</sup> British Geological Survey, Wallingford OX10 8BB, UK

## ARTICLE INFO

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Wolfgang Nowak, Associate Editor

## Keywords:

Groundwater drought  
Standardised Groundwater level Index (SGI)  
Spatio-temporal variability of groundwater levels  
Impulse response function  
Mixed model  
Kriging

## ABSTRACT

An empirical (geo)statistical modelling scheme is developed to address the challenges of modelling the severity and distribution of groundwater droughts given their spatially and temporally heterogeneous nature and given typically highly irregular groundwater level observations in space and time. The scheme is tested using GWL measurements from 948 observation boreholes across the Chalk aquifer (UK) to estimate monthly groundwater drought status from 1960 to 2013. For each borehole, monthly GWLs are simulated using empirical mixed models where the fixed effects are based on applying an impulse response function to the local monthly precipitation. Modelled GWLs are standardised using the Standardised Groundwater Index (SGI) and the monthly SGI values interpolated across the aquifer to produce spatially distributed monthly maps of SGI drought status for 54 years for the Chalk. The mixed models include fewer parameters than comparable lumped parameter groundwater models while explaining a similar proportion (more than 65%) of GWL variation. In addition, the empirical modelling approach enables confidence bounds on the predicted GWLs and SGI values to be estimated without the need for prior information about catchment or aquifer parameters. The results of the modelling scheme are illustrated for three major episodes of multi-annual drought (1975–1976; 1988–1992; 2011–2012). They agree with previous documented analyses of the groundwater droughts while providing for the first time a systematic, spatially coherent characterisation of the events. The scheme is amenable to use in near real time to provide short term forecasts of groundwater drought status if suitable forecasts of the driving meteorology are available.

## 1. Introduction

Groundwater drought is a type of hydrological drought defined as a period of below-normal groundwater level or reduced spring discharge (Tallaksen and Van Lanen 2004; Mishra and Singh, 2010; Van Loon, 2015). It is a natural hazard that can have a wide range of often profound social, economic and environmental impacts (Shahid and Hazarika 2010; Hughes et al., 2012; van Dijk et al., 2013; Medellín-Azuara et al., 2015) including: reductions in deployable output from boreholes potentially leading to costly restrictions to public supply; reduced abstractions for agriculture and industry; and, reduction in groundwater discharge to groundwater-supported rivers and wetlands which can cause ecological stress and lead to loss of amenity value (Lange et al., 2017). Consequently, there is a need to assess the severity and distribution of groundwater droughts to improve groundwater drought warnings; to contribute to improved water resource decision making during episodes of drought; and, to improve longer-term drought and water resources management plans (Van Lanen et al., 2016). However, assessment of the severity and distribution of

groundwater droughts is challenging due to a combination of issues primarily related to the availability and nature of groundwater level (GWL) data (Bachmair et al., 2016; Van Loon et al., 2017), and because groundwater droughts are commonly spatially and temporally heterogeneous in character (Peters et al., 2006; Mendicino et al., 2008; Tallaksen et al., 2009; Bloomfield et al., 2015).

Although there is increasing use of telemetry, GWLs are still often measured at relatively low frequencies, on the order of days to months or seasons, depending on the purpose of the groundwater monitoring (Li et al., 2015). In addition, GWL data are often temporally irregular and may contain artefacts. As a result, the availability of suitable GWL data at timescales relevant to the systematic assessment of droughts is often limited (Van Loon et al., 2017). Hydrological droughts propagate from a spatially and temporally heterogeneous meteorological signal (precipitation deficit), through soils (soil moisture deficit), and, via reduced recharge, lead to reduced groundwater levels (Peters et al., 2003). The spatio-temporal characteristics of the drought signal changes as it passes through the terrestrial water cycle so that groundwater droughts are typically lagged and attenuated compared

\* Corresponding author.

E-mail address: [benmarch@bgs.ac.uk](mailto:benmarch@bgs.ac.uk) (B.P. Marchant).

with meteorological droughts (Van Loon, 2015). There are many catchment and aquifer-related factors that influence these changes in drought signal and that can affect the eventual spatio-temporal expression of groundwater droughts including: the nature of land cover and the thickness and hydraulic characteristics of the soil and unsaturated zone that influence recharge; the hydraulic properties of the saturated zone; and, the location and hydraulic characteristics of major groundwater discharge areas (Tallaksen et al., 2009; Bloomfield et al., 2015).

This paper describes the first modelling method to take account of both the irregular and uncertain nature of groundwater level data and the spatial and temporally varying nature of droughts as they propagate through the terrestrial water cycle to estimate the monthly status of groundwater droughts including confidence bands on those estimates. The novel modelling scheme is tested with a large GWL data set from the Cretaceous Chalk limestone, the main aquifer in the United Kingdom. The results of the modelling are illustrated with the analysis of three major episodes of multi-annual drought across the Chalk (1995–1996; 1988–1992 and 2011–2012) and it is shown that they provide a high level of detail with regard to spatio-temporal variability of individual episodes of groundwater drought across the study area. Finally, the underlying causes of heterogeneity in the modelled GWLs and droughts are discussed, as are the wider applications of the modelling scheme for groundwater drought and groundwater resource assessments.

## 2. Modelling context

A range of methods have previously been used to model groundwater drought status, including: simple graphical representations of groundwater levels, for example of ranked historic records; the use of proxies of groundwater drought status, primarily to address issues related to the lack of suitable GWL data; process-based, distributed groundwater models; and, more recently, the use of multiple site specific, empirical and lumped parameter models.

Estimating groundwater drought status across a region can be as simple as relating recent observations of GWLs to historic data for a given site, for example by rank or percentile (Marsh et al., 2013; Natural Environment Research Council, 2017; USGS, 2017). However, such approaches rarely reveal sufficient detail of heterogeneity within episodes of groundwater drought to help with management responses (Van Loon et al., 2017). Bachmair et al. (2016) note in a recent survey of drought monitoring that there is ‘a lack of widespread monitoring of ... groundwater drought’ and that ‘the scarcity of water status observations, especially for groundwater, reflects the common focus on drought seen through the lens of rainfall and soil moisture that can be easily (remotely) monitored and/or modelled’. In this context, attempts have been made to overcome the common limitations of GWL data for drought status assessments by modelling GWL data with other more spatially extensive and/or higher frequency and/or more temporally regular hydrometric data, such as remotely sensed data, including, for example, data from the Gravity Recovery and Climate Experiment (GRACE) (Tapley et al., 2004). The modelled GWL data can then be used as proxies for groundwater drought status and hence used to obtain more detailed spatio-temporal descriptions of episodes of groundwater drought. However, remotely sensed data products can have low spatial resolution (e.g. 400 km grid squares for GRACE) and the use of remotely sensed data can be particularly sensitive to model parameterisation, such as depth to water table and depth to bedrock, and hence to spatial variations in often uncertain catchment properties (e.g. Li and Rodell, 2015).

Another approach to assess and analyse groundwater droughts is to use measures of precipitation deficits, such as the Standardised Precipitation Index (SPI) (McKee et al., 1993) or the Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano et al., 2010), as a proxy for groundwater drought and using linear models to

relate precipitation and GWLs (Khan et al., 2008; Bloomfield and Marchant, 2013; Bloomfield et al., 2015; Kumar et al., 2016; Van Loon et al., 2017). However, the correlation between SPI or SPEI and SGI is typically highly variable. One source of uncertainty in such simple correlations is because the effect of precipitation on GWLs varies over time, due, for example, to non-linear changes in soil moisture and degree of saturation in the unsaturated zone, or due to changes in saturated groundwater flow and discharge as a function of GWL. Consequently, a simple linear model relating precipitation and GWLs may not be appropriate. Impulse response functions (IRFs) (Von Asmuth et al., 2002; von Asmuth and Bierkens, 2005) can be used to describe the increase in GWL that will result from a single unit of precipitation as a function of time since the precipitation occurred. IRFs not only account for non-linear relationships between precipitation and GWLs, but also when estimated on a site-by-site basis will take account of spatial variations in such relationships, regardless of the specific non-linear recharge or discharge processes in operation in a catchment. IRFs have previously been used to describe relationships between effective precipitation and GWLs (Calver, 1997; Von Asmuth et al., 2002; von Asmuth and Bierkens, 2005); karst hydrographs and precipitation (Long and Derickson, 1999); stream-aquifer interactions (Hantush, 2005); and, spatial variation in recharge (Bakker et al., 2007).

Process-based, distributed groundwater models are used to undertake groundwater resource assessments at the catchment scale. They are designed to represent spatial variations in catchment and aquifer characteristics, and can be used to model groundwater extremes (Upton and Jackson, 2011). However, they require extensive data for calibration and can be difficult and costly to update frequently, so are of limited use in the assessment of the status of groundwater drought. Site specific, lumped parameter models are more flexible and typically easier to update and are increasingly being developed to provide a range of groundwater hydrological services, including GWL status assessments. For example, Aquimod, a lumped parameter model (Mackay et al., 2014; Jackson et al., 2016; Marchant et al., 2016) driven by rainfall data, is used to produce 1–3 month outlooks for GWL for the UK (Hydrological Outlooks, 2017; Prudhomme et al., 2017). Recently, Marchant et al. (2016) incorporated Aquimod as a fixed effect into a mixed model for simulation of GWLs and used it to quantify uncertainties in GWL simulations (GWL reconstructions) using formal likelihood methods. The mixed model approach enabled outliers and anomalies in the GWL data to be identified, periods of missing data to be infilled and for hydrographs to be reconstructed on regular dates within a formal uncertainty framework. However, even relatively simple lumped parameter models require a number of parameters, in the case of Aquimod – thirteen (Mackay et al., 2014), to either be fitted to the data or determined from other sources.

Building on the mixed model approach of Marchant et al. (2016), this paper reports on a simple approach to model groundwater drought status that both accounts for issues related to the availability and nature of GWL data and that takes account of catchment and aquifer heterogeneity and the heterogeneous propagation of droughts. An IRF, similar to that of Von Asmuth et al. (2002), is used as the fixed effect within the mixed model instead of a lumped parameter model, such as Aquimod. This results in a more parsimonious model with fewer parameters that need to be fitted (compared with Aquimod) while still accounting for the potentially spatially varying non-linear relationship between precipitation and GWL. The approach not only enables the issues of data quality to be addressed (for example, to screen out poor data or sites, infill missing data, and to model levels and drought status on a regular time step) but also enables uncertainty in the modelled levels to be quantified. The model does not require parameterisation using information on site characteristics and hence there are no *a priori* assumptions necessary about spatial variation in catchment or hydrogeological characteristics.



Contains Ordnance Data © Crown Copyright and database rights 2018. Ordnance Survey Licence no. 100021290

Fig. 1. Location and primary features of the Chalk Aquifer in the UK (left). Sites of the 948 boreholes used in this study with boreholes 'A' and 'B' highlighted (right).

### 3. Study area and data

#### 3.1. The study area

Fig. 1 shows the location of major features of the study area, the extent of the Chalk aquifer in England and overlying Clay-with-Flints and Crag deposits (Allen et al., 1997) and the location of observation boreholes used in the spatio-temporal analysis of groundwater drought status. The Chalk extends from the Yorkshire coast in the north-east of England, southwards to East Anglia, along the Chilterns and Berkshire Downs and throughout the London Basin, to the North and South Downs on the margins of the Weald, and westward on the margins of and beneath the Hampshire (Wessex) Basin. Although mapped as a single geological unit, the Chalk exhibits regional variations in stratigraphy and geological structure and intrinsic hydrogeological characteristics (such as porosity, permeability, transmissivity, storage and hydraulic diffusivity) over a range of scales (Downing et al., 2005; Allen et al., 1997).

The Chalk aquifer is a dual porosity, dual permeability limestone aquifer, where groundwater storage and flow are primarily dependent on the nature of the fracture system which may locally be sub-karstic (Bloomfield, 1996; Allen et al., 1997; MacDonald and Allen, 2001; Downing et al., 2005; Maurice et al., 2006). Groundwater levels in the Chalk respond to local hydro-climatology, they are sensitive to episodes of drought and depend on a combination of catchment characteristics, such as unsaturated zone thickness, and intrinsic hydraulic properties of the aquifer, such as hydraulic diffusivity (Bloomfield and Marchant, 2013). A characteristic feature of the Chalk aquifer, as noted by Bloomfield et al. (2015) and Rudd et al. (2017), is that, compared with other aquifers in the UK, such as the Permo-Triassic Sandstones and Jurassic Limestone aquifers, it is relatively susceptible to droughts, typically experiencing more severe and prolonged responses to droughts than other major aquifers in the UK.

As a first-order approximation, the broad meteorological drought history of the study area can be considered to be spatially homogeneous. This claim is consistent with the previously documented spatial coherence of major hydrological droughts in the UK (see “region 4” of Hannaford et al., 2011). For example, during the period 1960–2013 a number of episodes of major multi-annual drought affected the Chalk aquifer across the UK, notably: 1975–1976; 1988–1992; 1995–1997; 2004–2006; and 2010–2012 (Marsh et al., 2007; Kendon et al., 2013; Bloomfield and Marchant, 2013; Bloomfield et al., 2015). However, within the study area, spatio-temporal variations in the nature of groundwater droughts may be expected due to both spatial variation in the driving meteorology and the effects of heterogeneity in catchment and aquifer characteristics variably modifying the meteorological signals as they pass through to the groundwater system (Van Loon, 2015).

#### 3.2. Data

The borehole locations and GWL observations were downloaded from the Environment Agency's GWL monitoring network database (Environment Agency, 2014). The database included more than 4500 sites. The sites are monitoring or observation boreholes and wells that are typically open throughout their length in the Chalk and are not pumped or used for groundwater abstraction. The termination depths and construction details for individual boreholes and wells were not available to this study and no screening of the data had been conducted prior to downloading of the data. However although the observation boreholes represent a wide range of hydrogeological settings, including: unconfined and confined Chalk; sites close to or more distant to recharge and discharge areas; and, observation boreholes that may be variably affected by groundwater withdrawals from neighbouring abstraction boreholes or the effects of long-term changes in water resource management, the majority of sites reflect groundwater level variations

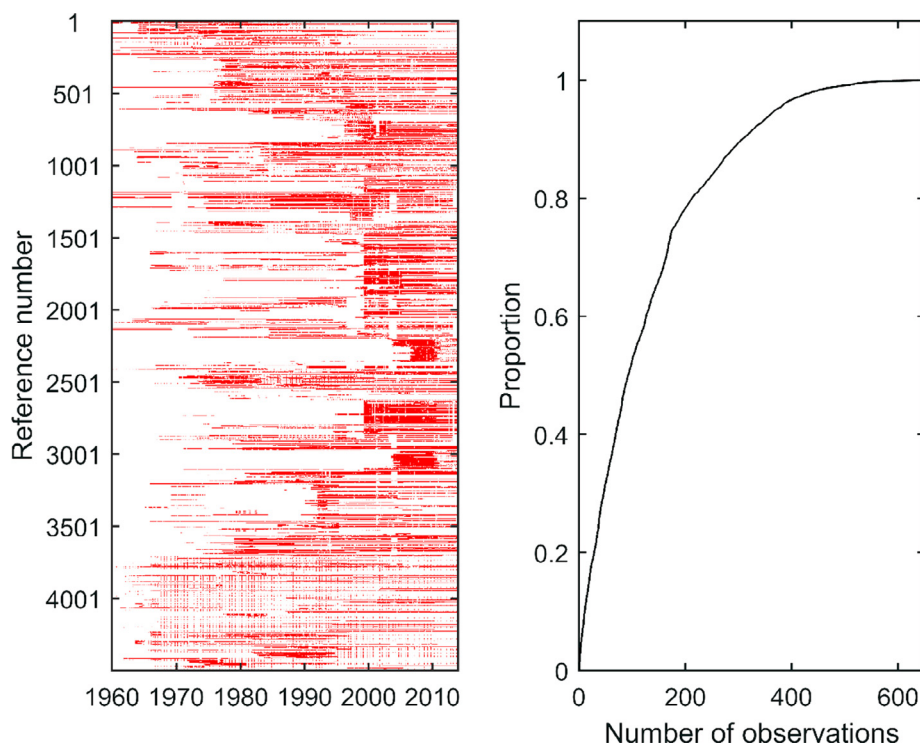


Fig. 2. Months for which GWL data is available for each borehole (left) and proportion of boreholes with specified number of monthly observations (right).

in the unconfined Chalk aquifer. There was also great variation in the temporal frequency of GWL observations. Some borehole records include periods of sampling every 15 min whereas others included fewer than 10 observations in total. The GWL data was subsampled so that there was no more than one observation per borehole per month. The resultant distribution of monthly observations is shown in Fig. 2. Fewer than 10% of the borehole records included an observation for more than half of the months between 1960 and 2013. Note also that there are some time periods, such as during the UK foot and mouth crisis in 2001, where very few observations are made because workers were restricted from visiting the boreholes (Fig. 2).

The precipitation data are taken from the CEH-GEAR dataset (Tanguy et al., 2015). This dataset consists of a 1 km<sup>2</sup> gridded estimate of monthly rainfall derived from the Met Office national database of observed precipitation. CEH-GEAR uses a natural neighbour interpolation methodology, including a normalisation step based on average annual rainfall, to generate the daily and monthly estimates. The total monthly precipitation from 1960 to 2013 has been extracted for each CEH-GEAR 1 km<sup>2</sup> grid square that contains a groundwater level observation borehole.

## 4. Methods

### 4.1. Overview of modelling workflow

The modelling workflow separately addresses the temporal (Fig. 3) and spatial (Fig. 4) variation of GWLs to produce spatiotemporal predictions of groundwater drought status. This approach contrasts with the use of space-time models (e.g. De Cesare et al., 2001; Li et al., 2016) to simultaneously account for both the spatial and temporal correlation amongst a set of observations of an environmental variable. Space-time models can be used to estimate a variable of interest at a particular location and time using data that was not observed at either the same time or the same location. However, they require the assumptions that the degree of spatial and temporal correlation exhibited by the data is constant across the study region and for all times. This assumption is inappropriate for GWL variation the Chalk where markedly different

patterns of temporal correlation are apparent at different boreholes.

The temporal portion (Fig. 3) of the workflow treats each borehole separately and uses the available GWL and precipitation data to produce 1000 simulations of the SGI on a monthly time step. The SGI simulations from each borehole are then combined in the spatial portion of the workflow (Fig. 4) which leads to (i) the identification of spatial clusters of boreholes which exhibit similar patterns of SGI variation and (ii) space-time predictions of the unimpacted drought status across the Chalk for a monthly time-step between 1960 and 2013.

### 4.2. Temporal modelling of groundwater levels

Temporal models of the GWLs are required in each borehole so that the GWLs might be predicted in months when they were not observed and to quantify the uncertainty of these predictions. Marchant et al. (2016) demonstrated that this could be achieved for long and irregularly sampled GWL records by coupling mechanistic models of GWLs, such as *AquiMod* (MacKay et al., 2014), with a statistical or empirical model of the residuals. They represented the observed GWLs for month  $m$ ,  $z(m)$ , by a mixed model:

$$z(m) = f(m) + r(m), \quad (1)$$

where  $f(m)$  was the estimated GWL according to *AquiMod* and  $r(m)$  was the model residual. The time series of residuals was assumed to be realized from a second order stationary random function with zero mean that was characterised by a marginal distribution function and a temporal correlation function. In the mixed model context the  $f(m)$  are referred to as the fixed effects and the  $r(m)$  as the random effects (Dobson, 1990).

However, for the majority of the boreholes in the Chalk there are insufficient observations to estimate the large number (at least 13) of *AquiMod* parameters. Instead, the approach based on IRFs proposed by von Asmuth et al. (2002) is used to relate the GWLs to the observed rainfall.

Von Asmuth et al. (2002) suggested that GWL response to rainfall could be modelled using a convolution between an IRF and the observed time series of effective rainfall i.e.:



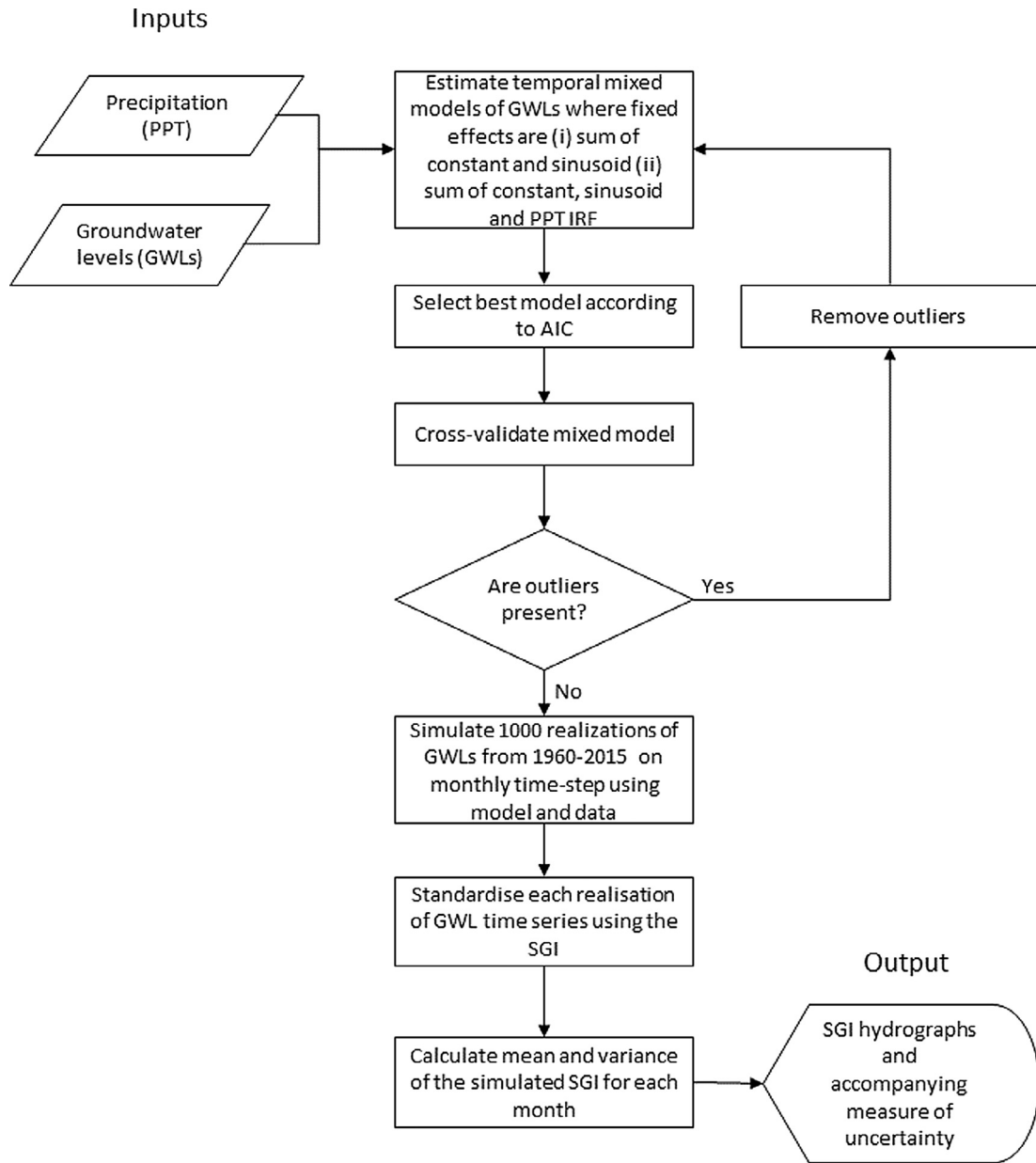


Fig. 3. Flow diagram of the temporal modelling of GWLs and SGIs for each borehole.

$$f(m) = \beta_0 + \int_0^m \alpha(\tau) p_e(m-\tau) d\tau, \quad (2)$$

where  $\alpha(\tau)$  is the impulse response function,  $p_e(m)$  the effective rainfall in month  $m$  and  $\beta_0$  is the GWL that occurs in the absence of any rainfall. The effective rainfall is the rainfall minus the water that is lost to evapotranspiration. Their IRF was based on a Pearson type III distribution:

$$\alpha(\tau) = A \frac{a^s \tau^{s-1} \exp(-a\tau)}{\Gamma(s)}, \quad (3)$$

where  $A$ ,  $a$  and  $s$  are parameters and  $\Gamma(s)$  is the gamma function of order  $s$ . If the effective rainfall is recorded every month then the integral in (2) may be discretized such that:

$$f(m) = \beta_0 + \sum_{\tau=0}^m \alpha(\tau) p_e(m-\tau). \quad (4)$$

In the UK, natural seasonal patterns of GWL variability are primarily the result of the seasonal variation in evapotranspiration. So here,

effective rainfall is replaced by the rainfall  $p(m)$  and sinusoids of periods of 6 and 12 months are added to the fixed effects to represent the seasonal variation caused by evapotranspiration. The summation in Eq. (4) is also restricted to the previous  $n_m = 60$  months. Hence, if  $\mathbf{m} = [m_1, m_2, \dots, m_n]^T$  is the vector of months for which GWLs were recorded then the fixed effects can be written:

$$f(\mathbf{m}) = \mathbf{X}\boldsymbol{\beta}, \quad (5)$$

where  $\mathbf{X}$  is the  $n \times 5 + n_m$  matrix:

$$\begin{bmatrix} 1 & \sin\left(\frac{2\pi m_1}{12}\right) & \cos\left(\frac{2\pi m_1}{12}\right) & \sin\left(\frac{2\pi m_1}{6}\right) & \cos\left(\frac{2\pi m_1}{6}\right) & p(m_1) & \cdots & p(m_1 - n_m + 1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \sin\left(\frac{2\pi m_n}{12}\right) & \cos\left(\frac{2\pi m_n}{12}\right) & \sin\left(\frac{2\pi m_n}{6}\right) & \cos\left(\frac{2\pi m_n}{6}\right) & p(m_n) & \cdots & p(m_n - n_m + 1) \end{bmatrix} \quad (6)$$

and  $\boldsymbol{\beta} = [\beta_0, \dots, \beta_{4+n_m}]^T$  is a vector of fixed effects parameters. The  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  and  $\beta_4$  are the parameters for the sinusoid terms, and the remaining  $\beta_{4+i} = \alpha(i)$ . It is assumed that the random effects are realized

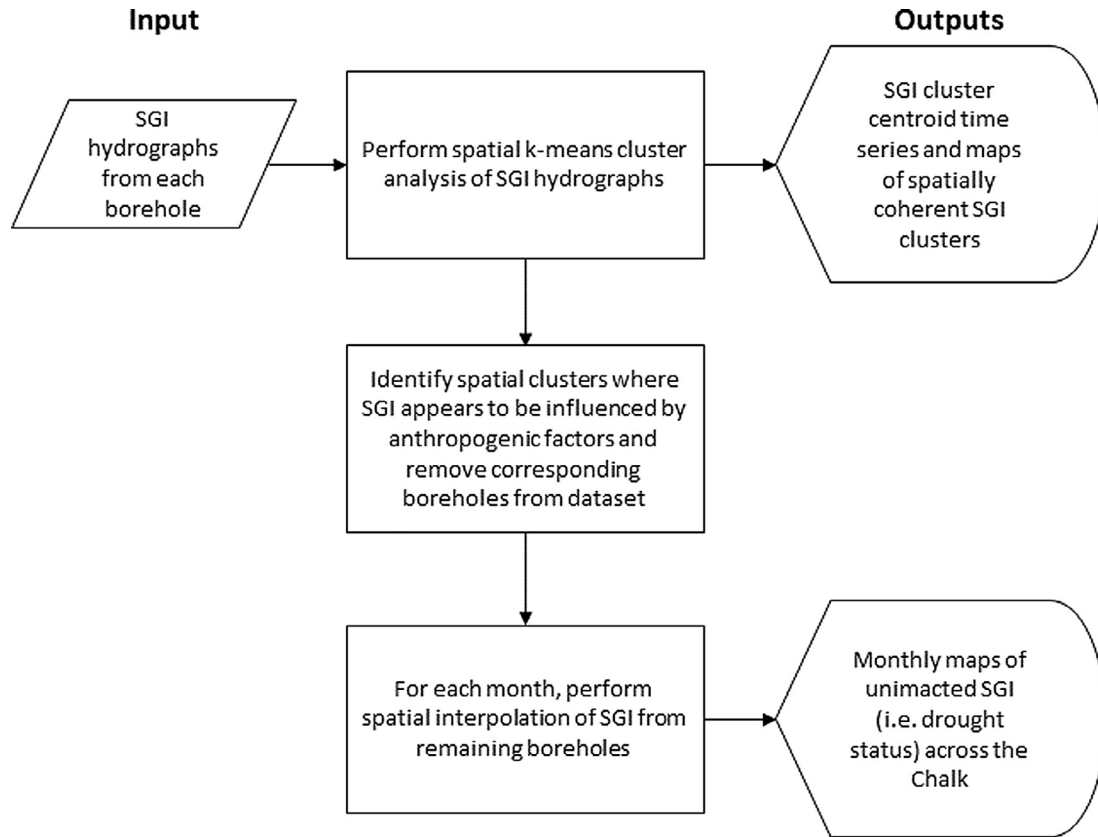


Fig. 4. Flow diagram of the spatial analysis of the modelled SGI from all of the boreholes.

from a second order stationary multivariate Gaussian random function where the correlation is represented by a nested nugget and Matérn function (Marchant and Lark, 2007):

$$C(h) = \begin{cases} c_0 + c_1 & \text{if } h = 0 \\ c_1 G(h) & \text{for } h > 0 \end{cases} \quad (7)$$

where

$$G(h) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \left( \frac{2\sqrt{\nu}h}{a} \right)^{\nu} K_{\nu} \left( \frac{2\sqrt{\nu}h}{a} \right). \quad (8)$$

Here,  $\Gamma$  is the Gamma function and  $K_{\nu}$  is a modified Bessel function of the second kind of order  $\nu$ . The random effects model parameters are  $c_0$  the nugget,  $c_1$  the partial sill,  $a$  the distance parameter and  $\nu$  the smoothness parameter.

Thus, the mixed model has a total of 12 parameters. Eight of these parameters are related to the fixed effects (five fewer than the minimum required by AquMod) and the remaining four are related to the random effects. The parameters values for each borehole are estimated by maximum likelihood (Marchant et al., 2016). The likelihood is a mathematical formula for the probability that a particular set of observations would have resulted from a specified statistical model. The maximum likelihood estimator uses a numerical optimization routine to find the parameter values which lead to the largest likelihood value for the observed data.

In some of the Chalk boreholes there was a negligible correlation between the fixed effects related to the rainfall IRF and the observed GWLs. This might have been because the variation of GWLs was primarily controlled by anthropogenic processes, such as abstraction, or because there were too few observations to reliably estimate the parameters of the IRF. In these circumstances the estimates of the IRF parameters are likely to be uncertain and the mixed model could be overfitted. This would mean that the model predictions on dates when the GWL was not observed are unreliable. To check for this occurrence a

simpler model of GWLs for each borehole is estimated where the fixed effects consisted of only a constant and the two sinusoids. The two models are compared by means of the Akaike Information Criterion (AIC; Akaike, 1973):

$$\text{AIC} = 2k - 2L, \quad (9)$$

where  $k$  is the number of model parameters and  $L$  is the logarithm of the maximised likelihood. The model that has the lowest AIC is selected. Thus, the AIC is a criterion by which it is tested whether the additional three parameters of the IRF lead to a substantial improvement of the fit of the model.

Once the best fitting mixed model has been estimated it is used along with the observed GWLs to predict GWLs on dates when they were not observed within the best linear unbiased predictor (Marchant et al., 2016) which is commonly referred to in the geostatistics literature as the universal kriging predictor (Webster and Oliver, 2007). This predictor yields a prediction,  $\hat{z}_i$ , of the expected GWL on each date  $t_i$  and a covariance matrix for these predictions. The entries on the main diagonal of this covariance matrix are the prediction variances,  $\sigma_i^2$ , which reflect the uncertainty of each individual prediction whereas the off-diagonal elements describe the relationship between the errors on different dates.

The prediction variances can be used to validate the estimated mixed model via the standardised squared prediction error (SSPE):

$$\theta_i = \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2},$$

where  $\hat{z}_i$  is the GWL predicted by the model for month  $i$ ,  $\sigma_i^2$  is the corresponding prediction variance and  $z_i$  is the observed GWL for month  $i$ . Observation  $z_i$  must be excluded from the predictor when calculating  $\hat{z}_i$ . A leave-one-out cross validation is employed, where each observation is omitted in turn and the remainder are used to predict the GWL on the corresponding date and to calculate the SSPE. If the errors

are normally distributed, then the SSPE should be realized from a standardised chi-squared distribution and the expected value of the mean of the SSPEs is equal to 1. The SSPE provides a means by which to identify the small number of typographical errors that are contained in the dataset. If a  $\hat{\theta}_i$  value is greater than a threshold set at 50 then this indicates that  $z_i$  is an outlier which might have undue influence on the model estimation procedure. In this case the outlier is removed from the dataset and the mixed model is re-estimated. A large threshold value for the outliers has intentionally been chosen (the expected number of false positives across the whole dataset according to the fitted model is fewer than  $10^{-6}$ ) to ensure that the removed observations are errors in data recording rather than examples of extreme GWLs.

Once the model has been estimated it is used in combination with the available data to produce uncertain simulations of the GWLs on a monthly time-step. The Cholesky simulation approach (Deutsch and Journel, 1998) is applied to the predicted covariance matrix,  $\mathbf{C}$ , to produce 1000 realizations of the errors in the predicted GWLs. The predictions are added to these simulations and the resultant time-series are de-seasonalized and normalised each of the resultant time to produce 1000 simulations for the borehole. De-seasonalization and normalisation is performed by applying the Standardised Groundwater Index (Bloomfield and Marchant, 2013), a non-parametric estimation that uses a normal-scores transform (Webster and Oliver, 2007) of groundwater level data for each calendar month. The nonparametric normalisation assigns a value to observations groundwater levels for a given month. The normal scores transform is undertaken by applying the inverse normal cumulative distribution function to  $n$  equally spaced  $p_i$  values ranging from  $1/(2n)$  to  $1-1/(2n)$ . The values that result are the SGI values for the given month. These are then re-ordered such that the largest SGI value is assigned to the  $i$  for which  $p_i$  is largest, the second largest SGI value is assigned to the  $i$  for which  $p_i$  is second largest, and so on. The normalisation is undertaken for each of the 12 calendar months separately and the resulting normalised monthly indices then merged to form a continuous SGI time series.

This modelling procedure was applied to the 948 boreholes with the most GWL observations. The remaining boreholes generally had too few observations to identify the parameters of the IRF.

#### 4.3. Cluster analysis of the SGI series

Groups of similarly behaving SGI time series are formed by performing a cluster analysis. Webster and Oliver (1990) discuss various clustering algorithms which lead to either hierarchical or non-hierarchical groups of individuals. Here, following Bloomfield et al. (2015), the flexible k-means clustering method is used to form a specified number of non-hierarchical groups. The k-means approach also requires that a distance function is specified which is used to assess the difference between different SGI time series. In this case, the Euclidean distance between the time series is used. A numerical optimisation routine is used to select the partitioning of the boreholes which minimizes the difference between each SGI time series and the centroid of the time series for the cluster in which it is contained. This procedure is repeated for two to nine groups and then expert interpretation is used to decide which number of groups is most suitable to reflect our understanding of the hydrogeological system.

The expert interpretation is guided by the following aims of the clustering: to identify, and so exclude from further analysis, sites that are systematically influenced by anthropogenic effects or that are outliers where the SGI time series do not correlate with the driving climatology; and that the remaining clusters are spatially-coherent and represent regional variations in groundwater droughts in the Chalk. The approach adopted is to identify the smallest number of clusters that reasonably satisfies the aims of the clustering. Note that these aims or rules are specific to the current study; however, for any given study area the target number of classes and hence the rules used can be adapted to reflect the regional hydrogeology and in particular any knowledge of

heterogeneity in the aquifer systems under investigation.

#### 4.4. Spatial interpolation of the SGI time series

Geostatistical methods (Webster and Oliver, 2007) are used to spatially interpolate simulated SGI values on each date. A spatial correlation function is estimated for the SGI values by the method of moments (Webster and Oliver, 2007). This approach calculates the semi-variance (i.e. half the squared difference) between pairs of observations and then estimates a model of how the semi-variance varies with the distance separating the boreholes at which the observations were made. In common with the temporal analyses the nested nugget and Matérn model (Eqs. (7) and (8)) is used. This ensures that the resultant correlation matrices are positive definite and that the prediction variances that result are not negative. When calculating the semi-variances comparisons are restricted to pairs of SGI values from the same date and from boreholes where the GWL was recorded on that date. This ensures that the temporal variation and the estimation variance of the SGI do not unduly influence our estimated semi-variances.

Having estimated the spatial correlation function, SGI values are spatially interpolated on the first day of each month, and the uncertainty of these predictions estimated by kriging (Webster and Oliver, 2007). The kriging predictor is applied 1000 times, taking a different simulated SGI value from each borehole each time. The SGI prediction at each unobserved site is equal to the mean prediction for these 1000 iterations. The variance of the 1000 predictions is calculated at each site and added to the spatial kriging variance to yield the total prediction variance. Thus this prediction variance accounts for both spatial and temporal uncertainty, although this approach does assume that the random effects from each borehole for a particular month are independent. Any location that is more than 75 km from a borehole is omitted from the region where predictions are calculated since the correlation between the GWLs at such a site and the observed GWLs will be small and hence the predictions unreliable.

### 5. Results

#### 5.1. Temporal modelling of SGI values

The results of the temporal modelling procedure are illustrated for two boreholes denoted 'A' and 'B' in Fig. 1 (right). Borehole 'A' is an observation borehole from the interfluvial area of the eastern Chilterns and has a monthly SGI hydrograph that shows annual and multi-annual responses to the driving meteorology and no clear annual minima, whereas borehole 'B' is an observation borehole from the northern part of the Hampshire Basin with a monthly SGI hydrograph that is dominated by an annual signal and a consistent annual minima, typically just below 80 m above sea level (masl), associated with a local groundwater base level. Fig. 5 shows the estimated IRFs for these two boreholes. The borehole 'A' IRF takes almost two years to reach its peak and precipitation is still having some influence on the GWLs 40 months after it has occurred. In contrast, the IRF for borehole 'B' reaches its peak after five months and it is negligibly small after 20 months. Thus, the modelled (and observed) GWLs from borehole 'A' vary relatively smoothly (Fig. 6 upper panel) whereas the GWLs from borehole 'B' (Fig. 6 lower panel) have a much flashier pattern of variation.

The 95% confidence limits in the predicted GWLs (Fig. 6) are determined from the values that are simulated from the estimated mixed model. The confidence limits for borehole 'A' are generally narrow although they widen during periods such as 1998 where few GWL observations are available. Similarly, the predictions of GWLs for borehole 'B' are rather uncertain in the early 1960s since this time precedes the observation of GWLs at this borehole. The SGI and 95% confidence limits of the SGI for each borehole are shown in Fig. 7. Again, these are based on the simulated GWLs and for borehole 'A' they are relatively narrow except at times when the GWLs were not observed. The

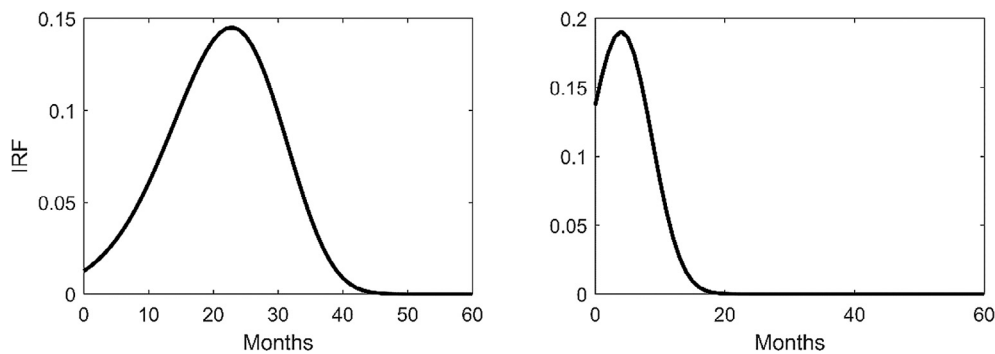


Fig. 5. Estimated impulse response functions for (left) Borehole 'A' and (right) Borehole 'B'.

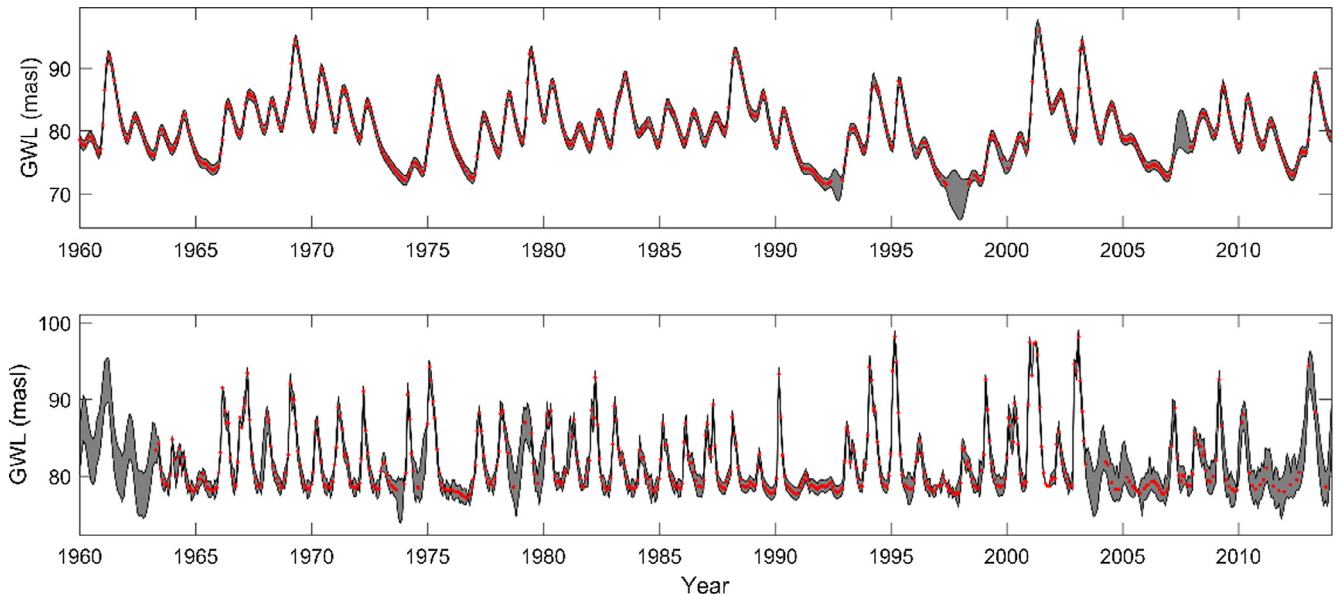


Fig. 6. Observed groundwater levels (red dots) and 95% confidence interval for predictions of groundwater levels (grey shaded area) for Borehole 'A' (upper) and Borehole 'B' (lower). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

confidence limits for the flashier SGI time series from borehole 'B' are much wider, particularly at times when the observed data were sparse.

When applying the temporal analysis workflow (Fig. 3) across all of the boreholes it was necessary to remove 28 outlying GWL observations which were inconsistent with the initial estimated model. Once these outliers had been removed and the models re-fitted the resultant mean SSPEs for each borehole were generally close to 1. For 93% of the boreholes (Table 1), the mean of the SSPE was between 0.9 and 1.1 indicating that the average observed squared errors upon cross-validation were on within 10% of the predicted squared errors. Of the best fitting models for each of the 948 boreholes analysed, 77% included an IRF and hence a significant relationship with precipitation (Table 1). Overall the best fitting models for all the sites explained an average of 52% of the variation in GWLs and on average more than half of this was explained by the IRF component.

## 5.2. Cluster analysis of SGI time series

Modelled GWLs from the 948 boreholes were used as the basis for the cluster analysis and the spatial distribution of sites for between two and nine clusters are shown in Fig. 8. Following the rules set for cluster selection, the results are shown in Figs. 8f and Fig. 9 and seven clusters have been identified. The centroids (i.e. averages) for the SGI time series for these seven clusters are shown in Fig. 10. The seven clusters include:

- two clusters of modelled hydrographs that show long-term trends and a poor explanation of observed variation in GWLs (Fig. 10 and Table 1) are interpreted as reflecting long-term anthropogenic influence on SGI time series, namely:
  - Cluster 7 (CL7) shows an upward trend from 1960 to 2013 (Fig. 10), where only 41% of the best fitting models include the IRF (Table 1) and only 21% of the GWL variation is explained by the model. Sites in this cluster are not spatially coherent and are interpreted as being affected by a range of anthropogenic influences, including, for example, previously documented post-industrial groundwater rebound in the confined Chalk beneath the central London (Lucas and Robinson, 1995) (Fig. 9c)
  - Cluster 3 (CL3) shows a downward trend (Fig. 10), where only 29% of the best fitting models include the IRF (Table 1) and only 12% of the GWL variation is explained by the model. Sites in this cluster, like those in CL7, are widely dispersed and are interpreted as being due to long-term or historic over abstraction from the Chalk, for example as previously documented in Lincolnshire (Whitehead and Lawrence, 2006) and Cambridgeshire (Petts et al., 1999; Acreman et al., 2000) (Fig. 9c).
- One cluster, Cluster 4 (CL4), of modelled hydrographs shows no trend and a very poor explanation of observed variation in GWLs (Table 1 and Fig. 9c). This cluster is interpreted as corresponding to sites where there is insufficient evidence (observations) to model the GWL variation.
- The remaining four clusters, Clusters 1, 2, 5 and 6 (CL1, CL2, CL5



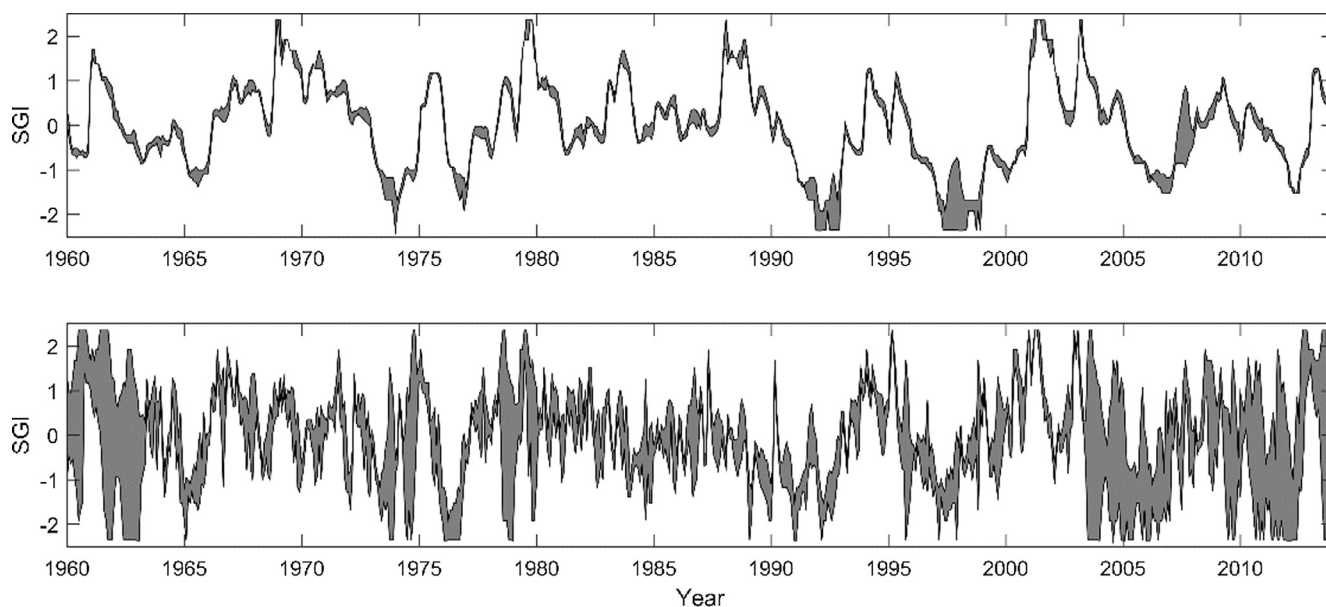


Fig. 7. Predicted 95% confidence intervals for SGI at Borehole 'A' (upper) and Borehole 'B' (lower).

Table 1

Number of boreholes in each cluster ( $n$ ); proportion of best fitting models including IRF (IRF model); proportion of variance of GWLs explained by model ( $R_{\text{tot}}^2$ ); proportion of variance of GWLs explained by IRF ( $R_{\text{IRF}}^2$ ) and proportion of boreholes where the mean SSPE is between 0.9 and 1.1 ( $0.9 < \bar{\theta} < 1.1$ ).

Cluster	$n$	IRF model	$R_{\text{tot}}^2$	$R_{\text{IRF}}^2$	$0.9 < \bar{\theta} < 1.1$
CL1	186	0.90	0.68	0.42	0.97
CL2	177	0.96	0.65	0.44	0.96
CL3	62	0.29	0.12	0.05	0.88
CL4	107	0.23	0.21	0.05	0.87
CL5	205	0.98	0.67	0.47	0.97
CL6	152	0.80	0.45	0.36	0.87
CL7	59	0.41	0.21	0.19	0.90
All	948	0.77	0.52	0.34	0.93

and CL6 respectively) occupy broadly spatially coherent regions, with CL1 coinciding with the Berkshire and North Wessex Downs, CL2 with Yorkshire, Lincolnshire, and north Norfolk, CL5 the South Downs and Wessex, and CL6 coincides with the Chilterns, North Downs and Suffolk. A high proportion of the models in each of these clusters include IRF models, and a high proportion of the variance of GWLs is explained by the models (Table 1).

Sites in clusters CL7, CL3 and CL4 are not analysed further since they are considered to be either anthropogenically impacted (CL7 and CL3) or the modelled GWLs are inferred to be unrelated to precipitation (CL4). The remaining four clusters, CL1, CL2, CL5 and CL6 are taken to represent the broadly unimpacted responses of groundwater levels in the unconfined or partially confined aquifer to the driving meteorology as modified by catchment and aquifer characteristics and are the focus of subsequent analysis.

The average (and standardised to a sum of one) IRFs for each of the four unimpacted clusters are shown in Fig. 11. The IRF for CL1 and CL2 are very similar, reaching a peak after about 7 months and becoming negligibly small after 40 months. Thus the SGI centroids for these two clusters (Fig. 10) have a similar degree of smoothness. The IRF for CL5 immediately reaches its peak and decreases to be negligibly small soon after 20 months. This behaviour is reflected by a much flashier SGI centroid in Fig. 10. Conversely, the IRF from CL6 remains positive beyond 50 months, corresponding to a very smooth SGI centroid. For the four unimpacted clusters, the vast majority of best fitting models

include the IRF (Table 1) and the models for CL1, CL2 and CL5 on average explain more than 65% of GWL variation. This level of performance is comparable with that achieved using Aquimod lumped parameter model (Jackson et al., 2016). On average only 45% of the variation in the GWLs is explained for CL6 boreholes, but this is consistent with the findings of Marchant et al. (2016) that predictive performance is relatively poor for smoothly varying GWLs. The mean SSPEs are generally close to one for all of the clusters (Table 1) although the proportion of sites where the mean SSPE is outside the range 0.9–1.1 is relatively large for the three impacted clusters and the more smoothly varying CL6.

To understand what may be influencing the composition of the clusters including their temporal and spatial features, it is helpful to consider both the precipitation time series associated with each SGI cluster and the spatial distribution of key catchment and hydro-geological factors.

Fig. 12 shows the mean normalised and standardised precipitation in the preceding 12 months,  $\text{SPI}_{12}$  (Bloomfield and Marchant, 2013), for clusters CL1, CL2, CL5 and CL6. Fig. 12 is consistent with the assumption that the broad drought history of the study area is spatially homogeneous. Time series of  $\text{SPI}_{12}$  for each of the four clusters closely parallel the  $\text{SPI}_{12}$  for the region as a whole, and all four show the previously documented major episodes of drought 1975–1976; 1988–1992; 1995–1997; 2004–2006; and 2010–2012 (Marsh et al., 1994; Marsh et al., 2007; Kendon et al., 2013; Bloomfield and Marchant, 2013; Bloomfield et al., 2015) (although it is noted that two additional episodes of  $\text{SPI}_{12}$  drought, taken here as  $\text{SPI}_{12} < -1$ , of a similar magnitude to previously recorded episodes of drought are also indicated for 1962–1966 and 1972–1973). There are, however, minor deviations in the  $\text{SPI}_{12}$  time series between individual clusters and the region as a whole. For example, the 1975–1976  $\text{SPI}_{12}$  drought was more acute in CL1, centred on the southern Chalk, rather than in CL2, i.e. over the Chalk of the north east; the 1988–1992 event appears to have been both more acute and continuous in CL2 than in CL1 and CL5 where it was slightly less severe and consisted of three individual phases; and, the onset of the 2011–2012 drought appears to be slightly earlier in CL1 than the other three clusters (Fig. 12). Even though episodes of groundwater drought, taken here as  $\text{SGI} < -1$ , are more attenuated than the equivalent drought in the driving meteorology (compare Fig. 10 with Fig. 12, and as would be expected from Van Loon, 2015), the subtle differences in the driving meteorology are

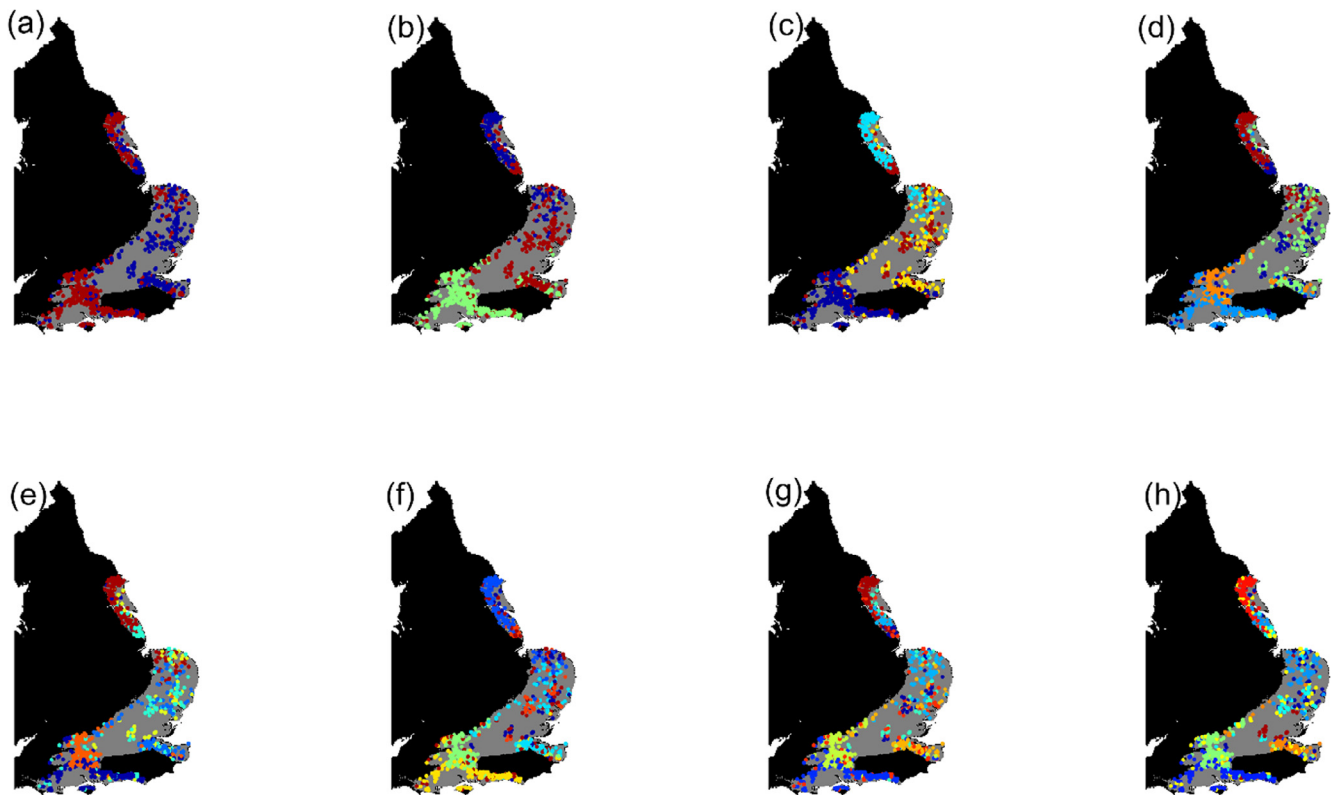


Fig. 8. Results of cluster analyses for two (a) to nine (h) clusters.

propagated through to and are reflected in similar differences in the SGI time series. For example, the SGI centroids for CL1 and CL2 are highly correlated (Fig. 10), but some differences are evident. The 1976 groundwater drought appears to be more severe for CL1, while the groundwater drought in the early 1990s is continuous for CL2 whereas in CL1 there is a small break, and the 2010–2012 groundwater drought commences slightly earlier in CL1. Based on the above, it appears that small changes in the driving meteorology within the study area, between CL2 in the north east of the study area and CL1 and CL5 in the south of the Chalk, have had some influence on the composition of the clusters.

The SGI centroids (Fig. 10) represent an integrated system response to the driving meteorology. However, there are features of the four clusters and their expression of groundwater droughts that appear to be

influenced by regional variation in unsaturated zone and recharge process and by some saturated zone and catchment characteristics.

Sites in CL6 may be influenced by features of the regional hydrogeology that act to modify recharge and hence affect the form of the SGI hydrographs. CL6 is associated with East Anglia, the Chilterns and the North Downs and is characterised by the smoothest of the SGI centroids (Fig. 10). Across large parts of East Anglia the Chalk is covered by the Crag, a series of younger gravels, sands and silts, and by glacial deposits (Jones et al., 2000). Weathering deposits, in the form of Clay-with-Flints, are also present extensively over the Chalk of the Chilterns and North Downs (Goudie, 1993; Bloomfield et al., 2009). The hydraulic relationship between these overlying deposits and the Chalk is locally complex. For example, the Crag may be in hydraulic continuity with the underlying Chalk or may be separated from the Chalk by a relatively

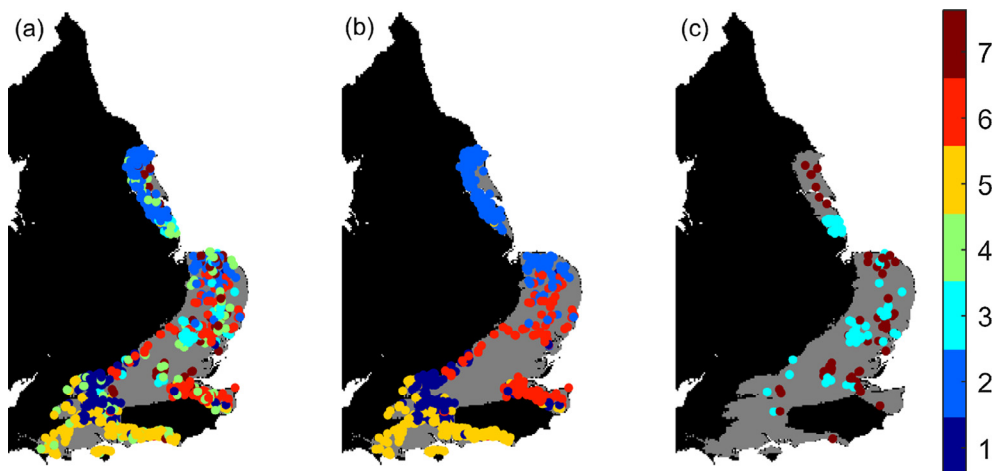


Fig. 9. Results of cluster analysis dividing the boreholes into seven clusters according to their SGI time series (a) all seven clusters, (b) the clusters where the SGI closely follows the precipitation time series (c) the clusters most impacted by abstraction.

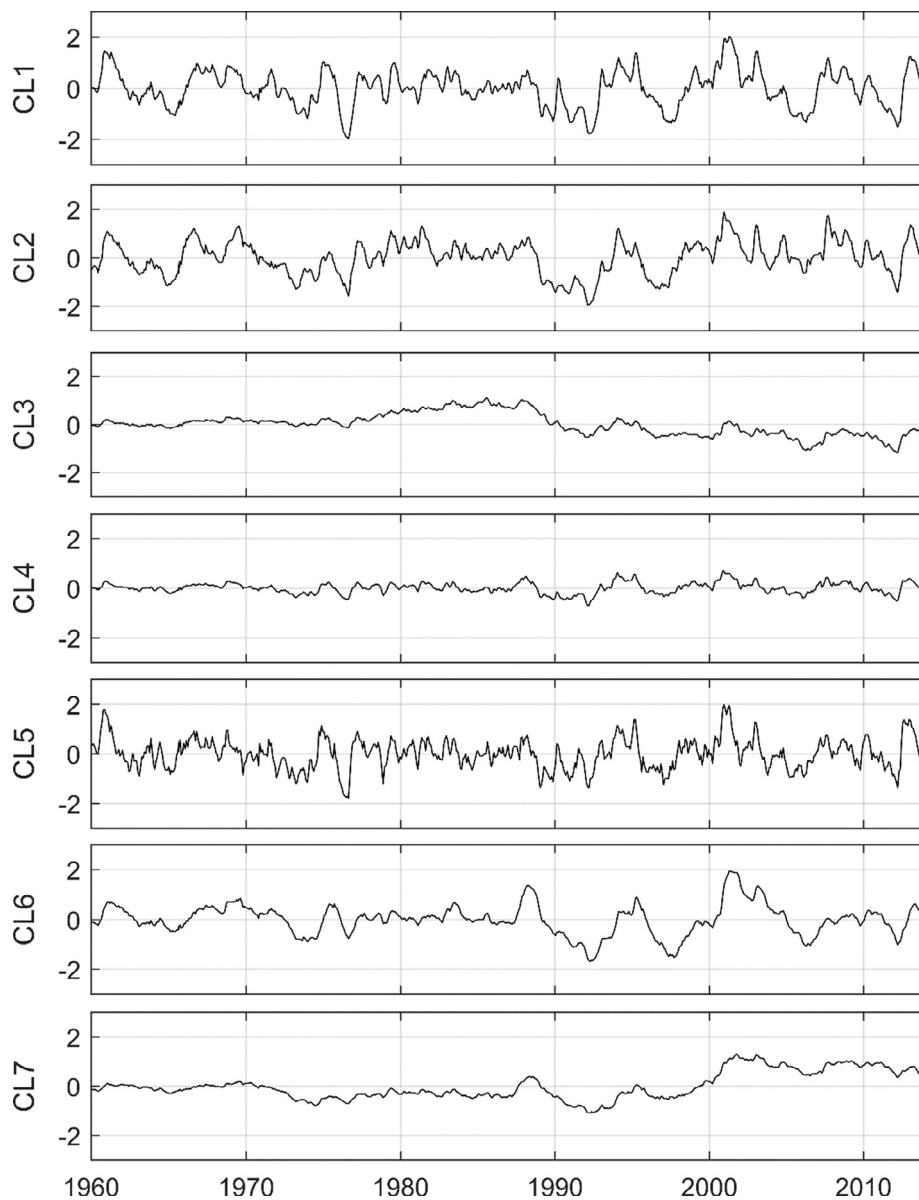


Fig. 10. The SGI centroids for each of the seven clusters shown in Fig. 9a.

low permeability clay (Jones et al., 2000). Where it is in hydraulic continuity it adds significantly to the storage of the aquifer, where the basal clay is present the Chalk may locally be confined. Over the Chilterns and the North Downs, the clay-with-flints deposits, another relatively low permeability clay, is typically found on high ground across the main recharge area associated with CL6. Here the clay-with-flints limits recharge, but also leads to focused recharge at the margins of the deposit (Allen et al., 1997; MacDonald et al., 1998). Although locally complex in their hydrogeological effects, at a regional scale these deposits that overly the Chalk may either add significant storage to the Chalk aquifer or limit or delay recharge. In both instances these processes will contribute to the relatively smooth nature of the SGI centroid of CL6 (Fig. 10).

Bloomfield and Marchant (2013) have previously shown that the range of temporal autocorrelation in the SGI from a borehole scales with the hydraulic diffusivity of the aquifer, consequently, it could be expected that the spatially coherent SGI clusters (Fig. 9b) should reflect any spatial heterogeneity in transmissivity (T) and storativity (S) of the Chalk at a regional scale. MacDonald and Allen (2001) analysed the aquifer properties of the Chalk based on 2100 pumping tests in four

regions: southern; the Thames Basin (including the North Downs); East Anglia; and Yorkshire and Lincolnshire, but the results were inconclusive. They found that the distribution of T and S was broadly similar for all four regions, and that within a given region transmissivity and storage co-efficients could vary by over five orders of magnitude. However, they also noted that the dataset was 'highly biased: most pumping tests have been undertaken in valley areas where the yield of the Chalk is highest', and it is unclear how this may have influenced the results of the analysis of the T and S data. The inference from these observations is that there appears to be no strong or reliable evidence for a relationship between T and S and the spatial distribution of the clusters. Notwithstanding this, relationships between the location of the observation borehole, catchment boundaries and local groundwater level base levels and hence unsaturated zone thickness will influence the form of the SGI time series (Peters et al., 2006; Bloomfield and Marchant, 2013). Compared with the other three clusters, catchments in Wessex and the South Downs associated with CL5 are typically relatively small and the Chalk is relatively highly faulted and fractured compared with other regions (Jones and Robins, 1999). These observations are consistent with the relatively flashy nature of the CL5 SGI

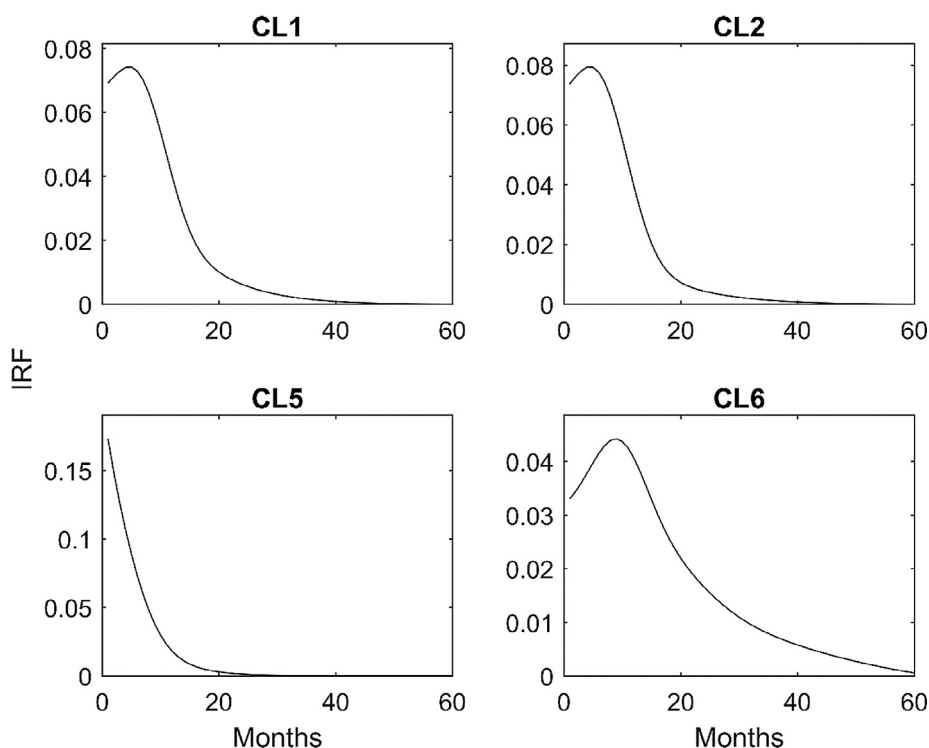


Fig. 11. The mean IRF for each of the four naturalised clusters shown in Fig. 9b.

centroid compared with CL1, CL2 and CL6 (Fig. 9).

In summary, it is not trivial to disentangle the various contributing factors to the form of the four SGI clusters. However, there is evidence

that small changes in the driving meteorology between the north east of the study area and the south of the study area may have influenced the formation of clusters CL1, CL2 and CL5; that overlying deposits and

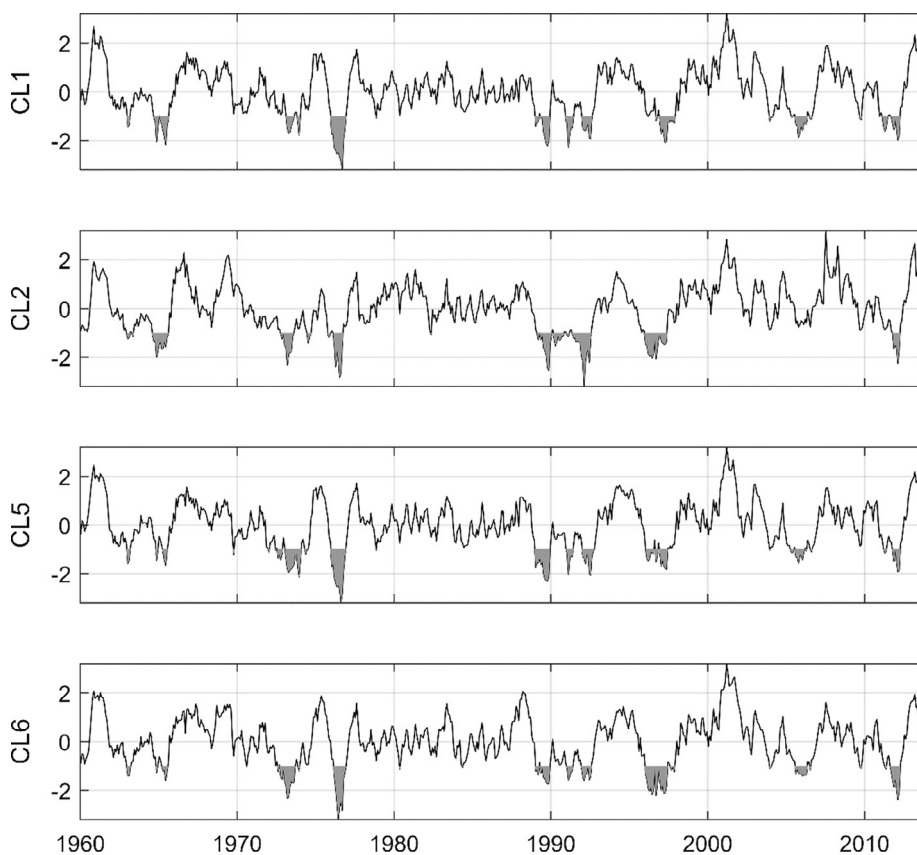


Fig. 12. Time series of normalised and standardised mean precipitation in the preceding 12 months for the four naturalised clusters shown in Fig. 9b. Shading regions correspond to SPI < -1.



their effect on unsaturated zone storage and recharge may have influenced the formation of CL6; and, CL5 is consistent with known spatial variations in catchment and rock mass characteristics.

### 5.3. Spatial interpolation of SGI values

The estimated spatial variogram function for the SGI predictions for all data in clusters CL1, CL2, CL5 and CL6 is shown in Fig. 13. The spatial variation is relatively smooth with little difference between SGI values a few km apart, and there is some degree of spatial correlation beyond 100 km. This is consistent with the spatial-coherency of the clusters, typically of the order of 100–200 km (Fig. 9b). Using the spatial variogram function, SGI values have been spatially interpolated for each month from January 1960 to March 2013.

Spatial interpolations of the SGI are shown for three major episodes of drought in Fig. 14 (1975–1975 event), Fig. 15 (1988–1992 event), and Fig. 16 (2011–2012 event) and reflect the observed differences between the SGI centroids for clusters CL1, CL2, CL5 and CL6 (Fig. 10). In Figs. 14–16 the distribution of negative values of SGI in the Chalk have been mapped where  $SGI < -0.5$  is denoted as pre-drought,  $< -1$  as groundwater drought,  $< -1.5$  as a severe groundwater drought and  $< -2$  as an extreme groundwater drought. Comparison of Figs. 14–16 shows the unique spatio-temporal nature of each episode of groundwater drought. For example, Fig. 14 shows that the 1975–1976 drought, although national in extent, at its peak in the summer of 1976 was most extreme in southern England (Hampshire and the Berkshire Downs), consistent with previous observations of Rodda and Marsh (2011). The 2011–2012 groundwater drought, Fig. 16, was also national in extent, however it was less pronounced than the 1975–1976 event, had a slightly more easterly footprint, and at its height, in late winter/spring 2012, was most extreme in north Norfolk and parts of the Chilterns and Berkshire, consistent with the observations of Marsh et al. (2013). Both the 1975–1976 and 2011–2012 episodes were effectively two-winter groundwater droughts where in any given region the drought status grew continuously over the entire period to a maximum and then declined as the drought broke (Marsh et al., 2013). However, in contrast to both these episodes, the 1988–1992 event was a longer, multi-annual event and is characterised by waves of increasing and decreasing groundwater drought status throughout the longer period. For example, Fig. 15 shows Yorkshire and parts of Hampshire passing into and back out of drought/severe drought status before finally reaching peak drought status in spring/summer 1992. Compared with both the 1975–1976 and 2011–2012 droughts, at its peak the 1988–1992 groundwater drought was more extensive across the whole of the English Chalk. However, both before and after peak drought

status (see March 1992, Fig. 15) the episode generally had a more north easterly focus than the other two episodes of drought (Marsh et al., 1994).

## 6. Discussion

The empirical approach adopted here successfully meets the challenges of assessing the severity and distribution of groundwater droughts while explicitly addressing issues associated with the availability and nature of GWL data and also accounting for and quantifying the spatio-temporal heterogeneity of episodes of drought. The empirical approach uses available GWL data for a region, in this case for the Chalk of the UK, and requires no additional information other than precipitation data. It does not need any process-based assumptions to be made and the models do not require parameterisation using information about site characteristics. As a consequence it is not necessary to have additional information about the spatial variation in catchment or hydrogeological characteristics over the modelled region. This means that the approach is applicable for use in monitoring and assessing groundwater drought status wherever there are suitable available GWL observations and precipitation data. The approach is more parsimonious with respect to the number of parameters that need to be calibrated than *AquiMod*, a process-based lumped model, while at the same time the performance of the empirical model has been shown to be comparable with that achieved using *AquiMod* (Jackson et al., 2016).

The modelling scheme (Figs. 3 and 4) takes GWL time series data, however irregular or noisy, and through sub-sampling and modelling transforms these into regular, monthly modelled GWL time series that can be correlated with precipitation data to estimate SGI and hence groundwater drought status. In this scheme, as with previous studies (Khan et al., 2008; Bloomfield and Marchant, 2013; Bloomfield et al., 2015; Kumar et al., 2016; Van Loon et al., 2017), precipitation is used as a proxy for groundwater drought status. However, the mixed model differs from and is an improvement on those previous studies that use proxies in two main ways. By using an IRF (Calver, 1997; Von Asmuth et al., 2002) the empirical approach accounts for non-linearities in the relationship between precipitation and groundwater levels. In addition, the method explicitly quantifies the uncertainty in both the modelled site and spatial groundwater drought status (Fig. 7). Fig. 17 shows spatial variation of the SGI prediction variance (i.e. uncertainty) across the Chalk for November 1962 and November 2012. The average prediction variances across the Chalk are 0.17 in November 1962 and 0.14 in November 2012. This difference reflects the greater number of groundwater observations in 2012. Spatially, the uncertainty increases with distance from the nearest borehole.

Van Lanen et al. (2016), Laaha et al. (2017), and Van loon et al. (2017) have each provided assessments of and commentaries on the last major European drought in 2015. Van Loon et al. (2017) note that there is a need ‘to promote more long-term groundwater measurement and international sharing of groundwater level data’. We re-emphasise this need. Data driven models of drought status, such as the one described here, require relatively long-term time series data across wide geographical areas to adequately represent variations in spatio-temporal status of groundwater droughts. Laaha et al. (2017) point out that a particular failing of hydrological assessments during the European drought of 2015 was the inability to obtain and analyse data in near real-time. In that context, any models of hydrological status that are calibrated on precipitation data (which is available in near real-time) and that can be used in a ‘nowcasting’ mode or for short-term forecasts of drought status (for example out to one or two months) would be highly valuable. Although outside the scope of the present study, the empirical model described here is amenable to be used in such a manner. The precision of these forecasts in a particular region will greatly depend upon the degree of temporal autocorrelation amongst the observed GWLs in nearby boreholes.

Van Lanen et al. (2016) noted that during droughts, such as that in

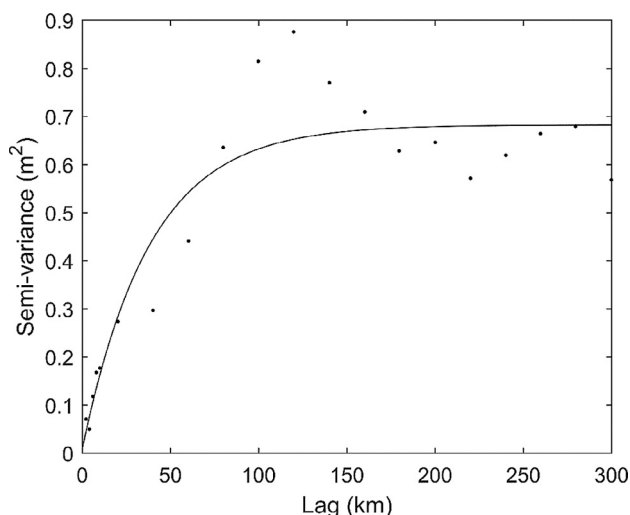


Fig. 13. Spatial variogram for observed SGI values.

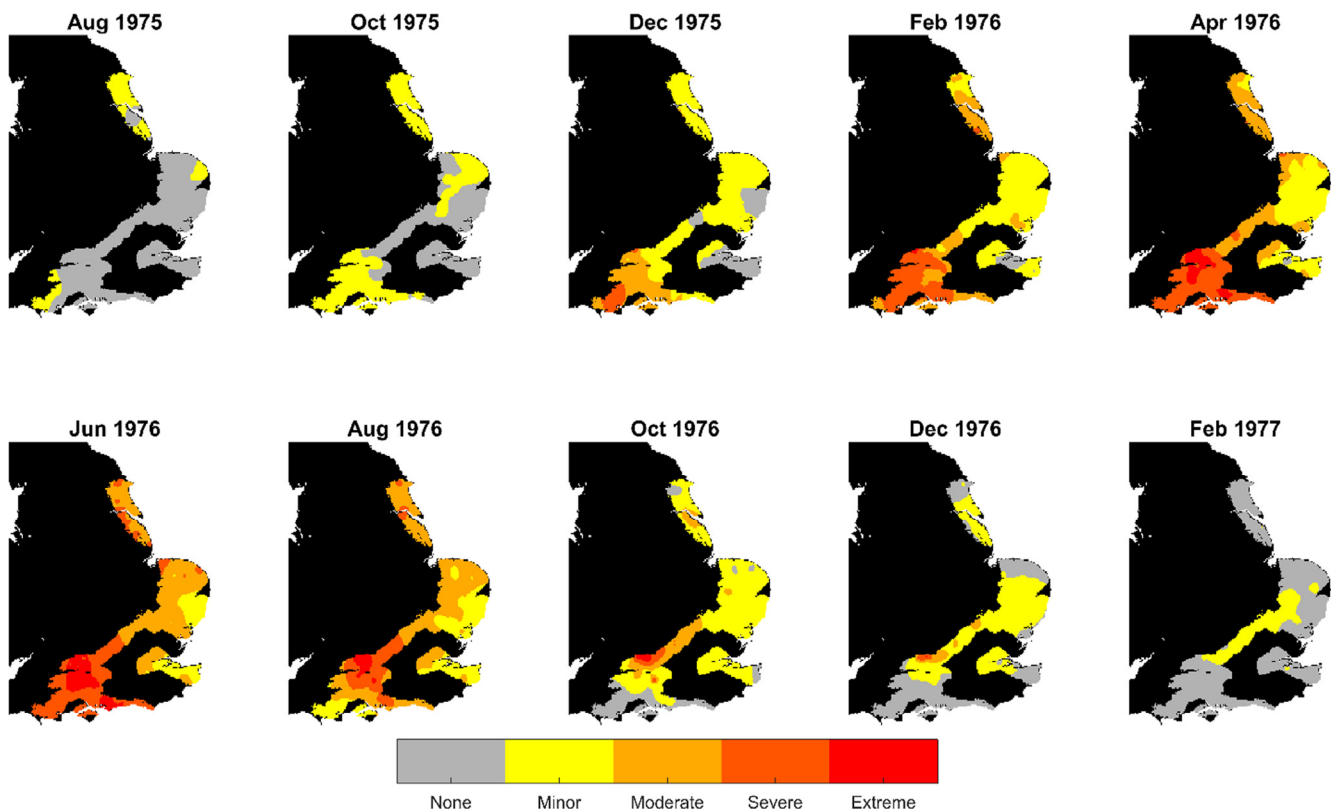


Fig. 14. Variation of groundwater drought status across the Chalk between August 1975 and February 1977. Note that the study area has been reduced to only include locations within 75 km of a borehole that is a member of CL1, CL2, CL5 or CL6.

Europe in 2015, groundwater resources are stressed not just due to precipitation deficits, but also due to over-abstraction and that typically ‘no separation is made between impacts due to the drought itself as compared to abstractions due to increased groundwater exploitation’. This observation has been further developed by [Van Loon et al. \(2016\)](#) in their analysis of drought in the Anthropocene. Although not the focus

of the current study, the modelling approach adopted here lends itself well to identification of the impacts of abstraction on groundwater droughts. Because the modelling does not rely on process-based assumptions of the hydrogeology at each site, it is easy to identify the degree to which variation in groundwater levels can be explained by variations in precipitation alone – where it is not possible for the model

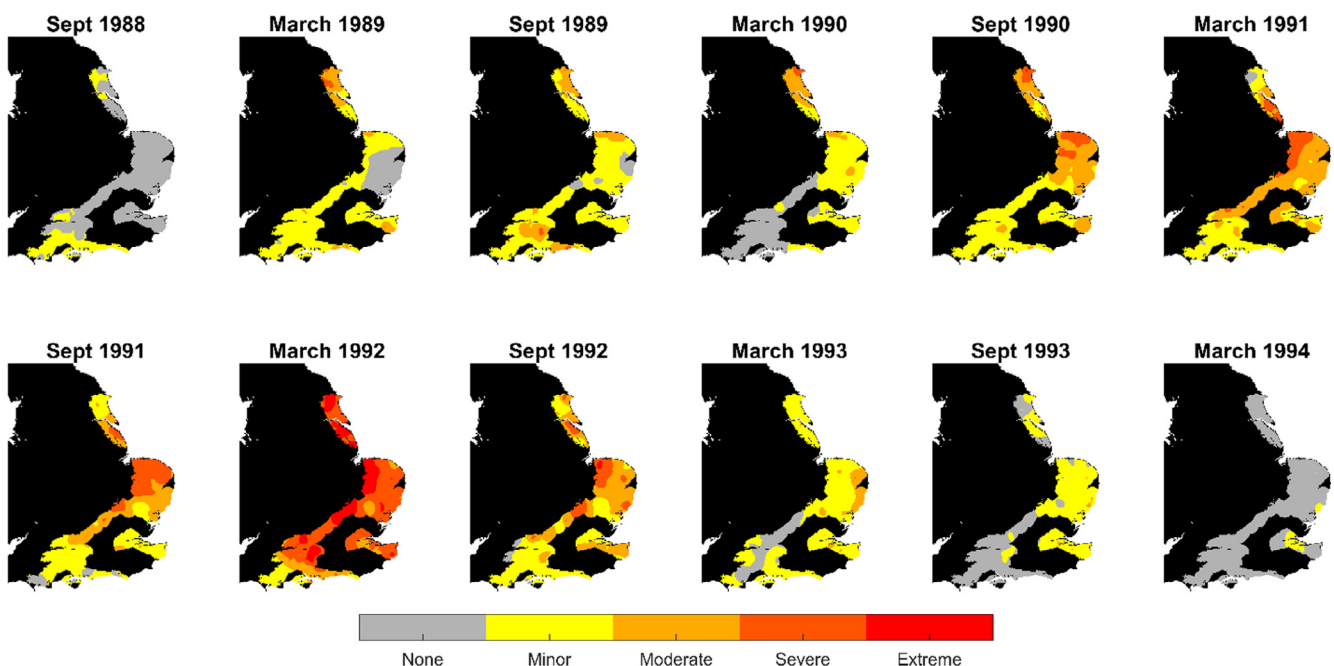
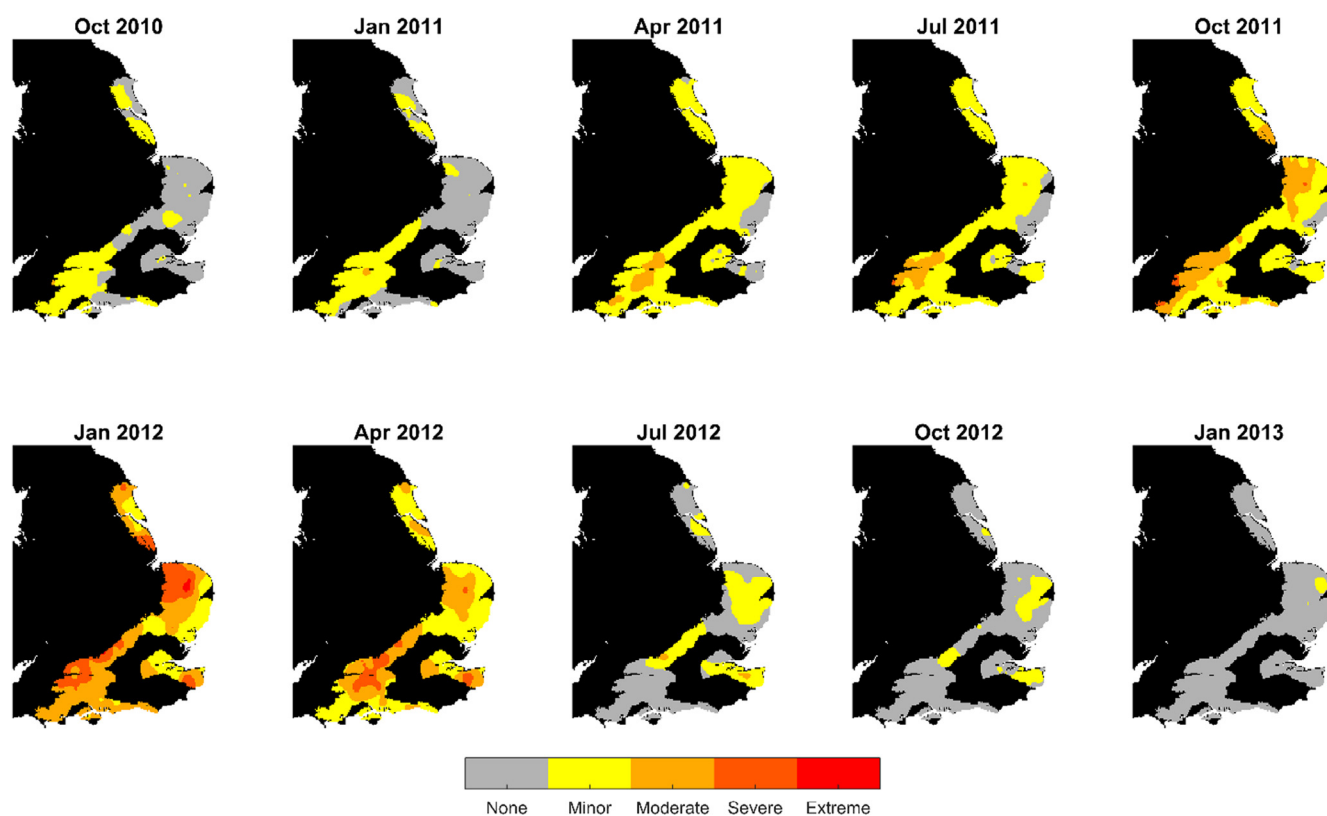


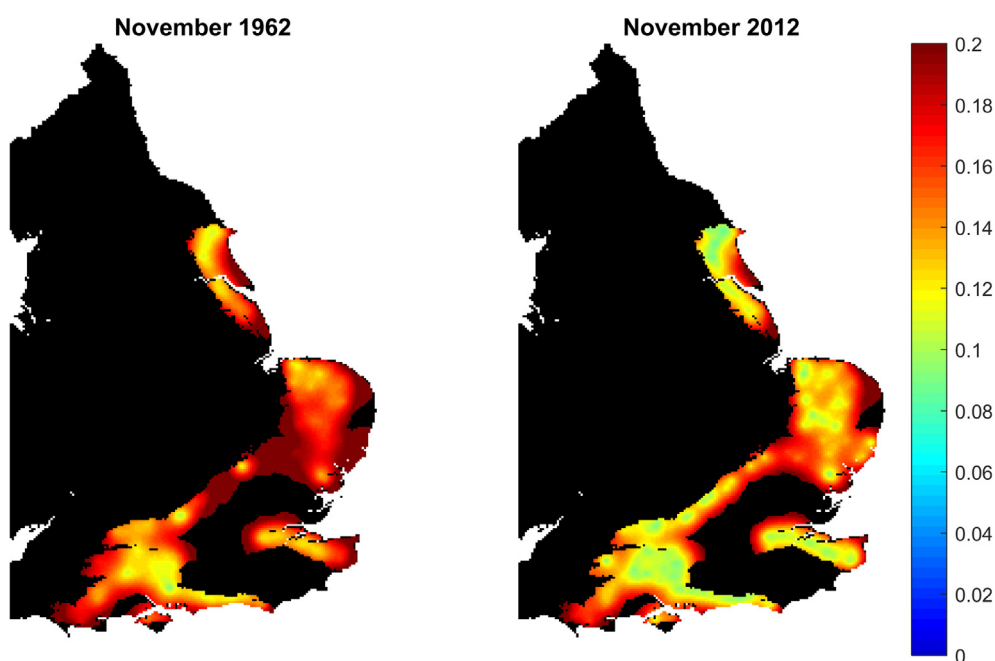
Fig. 15. Variation of groundwater drought status across the Chalk between September 1988 and March 1994. Note that the study area has been reduced to only include locations within 75 km of a borehole that is a member of CL1, CL2, CL5 or CL6.



**Fig. 16.** Variation of groundwater drought status across the Chalk between October 2010 and January 2013. Note that the study area has been reduced to only include locations within 75 km of a borehole that is a member of CL1, CL2, CL5 or CL6.

to explain groundwater levels adequately, it is reasonable to postulate that in many cases the effects of abstraction may be significant at that site. For example, in the present study, sites in clusters CL3 and CL7 have been inferred to be affected by long-term changes in groundwater management, with cluster CL3 interpreted as showing evidence of long-term abstraction. Extracting episodes of drought from hydrographs from CL3 where historic over abstraction has previously been

documented (e.g. [Whitehead and Lawrence, 2006](#)) and systematically analysing these in the context of equivalent hydrographs from cluster CL2 (hydrographs from the same region but that show no influence of long-term abstraction, [Figs. 9 and 10](#)) should provide new insights into the effects of long-term abstraction on the onset, magnitude and duration groundwater drought.



**Fig. 17.** Prediction variance for predicted SGI in November 1962 (left) and November 2012 (right).

## Acknowledgements

We would like to thank Mark Whiteman as well as staff at the Crommarsh Gifford office of the Environment Agency for facilitating access to the groundwater level data and we are grateful to Henry Holbrook for his work in the production of Fig. 1. The work has been funded by Natural Environment Research Council through the Historic Droughts project (grant number NE/L010151/1). This paper is published with the permission of the Executive Director of the British Geological Survey (Natural Environment Research Council).

## References

- Acreman, M.C., Adams, B., Birchall, P., Connorton, B., 2000. Does groundwater abstraction cause degradation of rivers and wetlands? *J. CIWEM* 14, 200–206.
- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petov, B.N., Csaki, F. (Eds.), *Second International Symposium on Information Theory*. Akademia Kiado, Budapest, pp. 267–281.
- Allen, D.J., Brewerton, L.J., Coleby, L.M., Gibbs, B.R., Lewis, M.A., MacDonald, A.M., Wagstaff, S.J., Williams, A.T., 1997. *The Physical Properties of Major Aquifers in England and Wales*, British Geological Survey Research Report WD/97/34. British Geological Survey, Keyworth, UK.
- Bachmair, S., Stahl, K., Collins, K., Hannaford, J., Acreman, M., Svoboda, M., Knutson, C., Smith, K.H., Wall, N., Fuchs, B., Crossman, N.D., Overton, I.C., 2016. Drought indicators revisited: the need for a wider consideration of environment and society. *Wiley Interdiscip. Rev. Water* 3, 516–536.
- Bakker, M., Maas, K., Schaars, F., von Asmuth, J.R., 2007. Analytic modeling of groundwater dynamics with an approximate impulse response function for areal recharge. *Adv. Water Resour.* 30, 493–504.
- Bloomfield, J.P., 1996. Characterization of hydrogeologically significant fracture distributions in the Chalk at Fair Cross, Berkshire, UK. *J. Hydrol.* 184, 355–379.
- Bloomfield, J.P., Marchant, B.P., 2013. Analysis of groundwater drought using a variant of the Standardised Precipitation Index. *Hydrol. Earth Syst. Sci.* 17, 4769–4787.
- Bloomfield, J.P., Allen, D.J., Griffiths, K.J., 2009. Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK. *J. Hydrol.* 373, 164–176.
- Bloomfield, J.P., Marchant, B.P., Bricker, S.H., Morgan, R.B., 2015. Regional analysis of groundwater droughts using hydrograph classification. *Hydrol. Earth Syst. Sci.* 19, 4327–4344.
- Calver, A., 1997. Recharge response functions. *Hydrology and Earth Systems Sciences* 1, 47–53.
- De Cesare, L., Myers, D.E., Posa, D., 2001. Estimating and modelling space-time correlation structures. *Statist. Probab. Lett.* 51, 9–14.
- Deutsch, C.V., Journel, A.G., 1998. *GSLIB Geostatistical Software Library and User's Guide*, second ed. Oxford University Press, New York.
- Dobson, A.J., 1990. *An Introduction to Generalized Linear Models*, second ed. Chapman and Hall, UK.
- Downing, R.A., Price, M., Jones, G.P., 2005. *The Hydrogeology of the Chalk of North-West Europe*. Clarendon Press, Oxford, UK, pp. 310.
- Environment Agency., 2014. *National groundwater level database for England*, Environment Agency. <http://data.gov.uk/data> last retrieved 26 August 2017.
- Goudie, A.S., 1993. *The Landforms of England and Wales*. revised ed., Wiley-Blackwell, pp. 408.
- Hannaford, J., Lloyd-Hughes, B., Keef, C., Parry, S., Prudhomme, C., 2011. Examining the large-scale spatial coherence of European drought using regional indicators of rainfall and streamflow deficit. *Hydrol. Process.* 25, 1146–1162.
- Hantush, M.M., 2005. Modeling stream-aquifer interactions with linear response functions. *J. Hydrol.* 311, 59–79.
- Hughes, J.D., Petrone, K.C., Silberste, R.C., 2012. Drought, groundwater storage and stream flow decline in southwestern Australia. *J. Geophys. Res. Lett.* 39, L03408.
- Jackson, C.R., Wang, L., Pachocka, M., Mackay, J.D., Bloomfield, J.P., 2016. Reconstruction of multi-decadal groundwater level time-series using a lumped conceptual model. *Hydrol. Process.* 30, 3107–3125.
- Jones, H.K., Robins, N.S., 1999. *The Chalk aquifer of the South Downs*. In: *Hydrogeological Report Series*, British Geological Survey. Natural Environment Research Council, Keyworth, UK.
- Jones, H.K., Morris, B.L., Cheney, C.S., Brewerton, L.J., Merrin, P.D., Lewis, M.A., MacDonald, A.M., Coleby, L.M., Talbot, J.C., McKenzie, A.A., Bird, M.J., Cunningham, J., Robinson, V.K., 2000. *The Physical Properties of Minor Aquifers in England and Wales*, British Geological Survey Technical Report WD/00/4. British Geological Survey, Keyworth, UK.
- Kendon, M., Marsh, T., Parry, S., 2013. The 2010–2012 drought in England and Wales. *Weather* 68, 88–95.
- Khan, S., Gabriel, H.F., Rana, T., 2008. Standard precipitation index to track drought and assess impact of rainfall on water tables in irrigation areas. *Irrig. Drain. Syst.* 22, 159–177.
- Kumar, R., Musuza, J.L., Van Loon, A.F., Teuling, A.J., Barthel, R., Ten Broek, J., Mai, J., Samaniego, L., Attinger, S., 2016. Multiscale evaluation of the Standardized precipitation Index as a groundwater drought indicator. *Hydrol. Earth Syst. Sci.* 20, 1117–1131.
- Laaha, G., Gauster, T., Tallaksen, L.M., Vidal, J.-P., Stahl, K., Prudhomme, C., Heudorfer, B., Vlnas, R., Ionita, M., Van Lanen, H.A.J., Adler, M.-J., Caillouet, L., Delus, C., Fendekova, M., Gailliez, S., Hannaford, J., Kingston, D., Van Loon, A.F., Mediero, L., Osuch, M., Romanowicz, R., Sauquet, E., Stagge, J.H., Wong, W.K., 2017. The European 2015 drought from a hydrological perspective. *Hydrol. Earth Syst. Sci.* 21, 3001–3024.
- Lange, B., Holman, I., Bloomfield, J.P., 2017. A framework for a joint hydro-meteorological-social analysis of drought. *Sci. Total Environ.* 578, 297–306.
- Li, H.Y., Marchant, B.P., Webster, R., 2016. Modelling the electrical conductivity of soil in the Yangtze delta in three dimensions. *Geoderma* 269, 119–125.
- Li, B., Rodell, M., Famiglietti, J.S., 2015. Groundwater variability across temporal and spatial scales in the central and northeastern U.S. *J. Hydrol.* 525, 769–780.
- Li, B., Rodell, M., 2015. Evaluation of a model-based groundwater drought indicator in the conterminous U.S. *J. Hydrol.* 526, 78–88.
- Long, A.J., Derickson, R.G., 1999. Linear systems analysis in a karst aquifer. *J. Hydrol.* 219, 206–217.
- Lucas, H.C., Robinson, V.K., 1995. Modelling of rising groundwater levels in the Chalk aquifer of the London Basin. *Q.J.E.G.H.* 28, S51–S62.
- MacDonald, A.M., Allen, D.J., 2001. Aquifer properties of the Chalk of England. *Q. J. Eng. Geol. Hydrogeol.* 34 (4), 371–384.
- MacDonald, A.M., Brewerton, L.J., Allen, D.J., 1998. Evidence for rapid groundwater flow and karst-type behaviour in the Chalk of southern England. In: Robins, N.S. (Ed.), *Groundwater Pollution, Aquifer Recharge and Vulnerability*. Geological Society, London.
- Mackay, J.D., Jackson, C.R., Wang, L., 2014. A lumped conceptual model to simulate groundwater level time-series. *Environ. Modell. Softw.* 61, 229–245.
- Marchant, B.P., Lark, R.M., 2007. The Matérn variogram model: Implications for uncertainty propagation and sampling in geostatistical surveys. *Geoderma* 140, 337–345.
- Marchant, B.P., Mackay, J., Bloomfield, J.P., 2016. Quantifying uncertainty in predictions of groundwater levels using formal likelihood methods. *J. Hydrol.* 540, 699–711.
- Marsh, T., Cole, G., Wilby, R., 2007. Major droughts in England and Wales, 1800–2006. *Weather* 62, 87–93.
- Marsh, T.J., Monkhouse, R.A., Arnell, N.W., Lees, M.L., Reynard, N.S., 1994. The 1988–92 drought. Wallingford, NERC Institute of Hydrology report.
- Marsh, T.J., Parry, S., Kendon, M.C., Hannaford, J., 2013. The 2010–12 Drought and Subsequent Extensive Flooding. Pub. Centre for Ecology & Hydrology.
- Maurice, L.D., Atkinson, T.C., Barker, J.A., Bloomfield, J.P., Farrant, A., Williams, A., 2006. Karstic behaviour of groundwater in the English Chalk. *J. Hydrol.* 330, 63–70.
- McKee, T. B., Doerken, N. J., Leist, J., 1993. The relationship of drought frequency and duration time scales, 8th Conference on Applied Climatology, 17–22 January 1993, Anaheim, California, 179–184, 1993.
- Medellin-Azuara, J., MacEwan, D., Howitt, R.E., Koruakos, G., Dogrul, E.C., Brush, C.F., Kadir, T.N., Harter, T., Melton, F., Lund, J.R., 2015. Hydro-economic analysis of groundwater pumping for irrigated agriculture in California's Central Valley, USA. *Hydrogeol. J.* 23, 1205–1216.
- Mendicino, G., Senatore, A., Versace, P., 2008. A Groundwater resource index (GRI) from drought monitoring and forecasting in a Mediterranean climate. *J. Hydrol.* 357, 282–302.
- Mishra, A.K., Singh, V.J., 2010. A review of drought concepts. *J. Hydrol.* 391, 202–216.
- Natural Environment Research Council, 2017. *Monthly Hydrological Summaries*. National River Flow Archive, UK.
- Hydrological Outlooks., 2017. *The Hydrological Outlook UK*. Web page published by Natural Environment Research Council: Centre for Ecology and Hydrology. <http://www.hydoutuk.net/> Last accessed 1 August 2017.
- Peters, E., Torfs, P.J.J.F., van Lanen, H.A.J., Bier, G., 2003. Propagation of drought through groundwater – A new approach using linear reservoir theory. *Hydrol. Process.* 16, 3023–3040.
- Peters, E., Bier, G., Van Lanen, H.A.J., Torfs, P.J.J.F., 2006. Propagation and spatial distribution of drought in a groundwater catchment. *J. Hydrol.* 321, 257–275.
- Petts, E., Bickerton, M.A., Crawford, C., Lerner, D.N., Evans, D., 1999. Flow management to sustain groundwater-dominated stream ecosystems. *Hydrol. Process.* 13, 497–513.
- Prudhomme, C., Hannaford, J., Boorman, D., Knight, J., Bell, V., Jackson, C., Svensson, C., Parry, S., Bachiller-Jareno, N., Davies, H.N., Davis, R., Harrigan, S., Mackay, J., McKenzie, A., Rudd, A.C., Smith, K., Bloomfield, J.P., Ward, R., Jenkins, A., 2017. *Hydrological Outlook UK: an operational streamflow and groundwater level forecasting system at monthly to seasonal time scales*. Hydrol. Sci.
- Rodda, J.C., Marsh, T.J., 2011. *The 1975–76 Drought – A contemporary and retrospective review*. Centre for ecology and Hydrology report, ISBN: 978-1-906698-24-9, 58p. <http://nora.nerc.ac.uk/15011/>
- Rudd, A., Bell, V.A., Kay, A.L., 2017. National-scale analysis of simulated hydrological droughts (1891–2015). *J. Hydrol.* 550, 368–385.
- Shahid, S., Hazarika, M.K., 2010. Groundwater drought in the northwestern districts of Bangladesh. *Water Resour. Manage.* 24, 1989–2006.
- Tallaksen, L.M., Hissdal, H., Van Lanen, H.A.J., 2009. Space-time modeling of catchment scale drought characteristics. *J. Hydrol.* 375, 363–372.
- Tallaksen, L.M., Van Lanen, H.A.J., 2004. *Hydrological drought: processes and estimation methods for streamflow and groundwater*. Developments in Water Science. Elsevier Science B.V.
- Tanguy, M., Dixon, H., Prosdocimi, I., Morris, D.G., Keller, V.D.J., 2015. Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890–2014) [CEH-GEAR]. NERC Environ. Inform. Data Centre. <https://doi.org/10.5285/f2856ee8-daf6-4b67-bdb-590520c77b3c>.
- Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F., Watkins, M.M., 2004. GRACE measurements of mass variability in the Earth system. *Science* 305 (5683), 503–505.
- Upton, K.A., Jackson, C.R., 2011. Simulation of the spatio-temporal extent of groundwater flooding using statistical methods of hydrograph classification and lumped parameter models. *Hydrol. Process.* 25, 1949–1963.



- USGS., 2017. USGS Groundwater Watch. Web page published by U.S. Department of the Interior/U.S. Geological Survey <https://groundwaterwatch.usgs.gov/> last accessed 1st March 2017.
- van Dijk, A.I.K.J.M., Beck, H.E., Crosbie, R.S., de Jeu, R.A.M., Liu, Y.Y., Podger, G.M., Timbal, B., Viney, N.R., 2013. The millennium drought in southeast Australia (2001–2009): natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resour. Res.* 49, 1040–1057.
- Van Lanen, H.A.J., Laaha, G., Kingston, D.G., Gauster, T., Ionita, M., Vidal, J.-P., Vlnas, R., Tallaksen, L.M., Stahl, K., Hannaford, J., Delus, C., Fendekova, M., Mediero, L., Prudhomme, C., Rets, E., Romanowicz, R.J., Gailliez, S., Wong, W.K., Adler, M.-J., Blauhut, V., Caillouet, L., Chelcea, S., Frolova, N., Gudmundsson, L., Hanel, M., Haslinger, K., Kireeva, M., Osuch, M., Sauquet, E., Stagge, J.H., Van Loon, A.F., 2016. Hydrology needed to manage droughts: the 2015 European Case. *Hydrol. Process.* 30, 3097–3104.
- Van Loon, A.F., 2015. Hydrological drought explained. *WIREs Water* 2, 359–392.
- Van Loon, A.F., Gleeson, T., Clark, J., Van Dijk, A.I.J.M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A.J., Tallaksen, L.M., Uijlenhoet, R., Hannah, D.H., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangelcroft, S., Wanders, N., Van Lanen, H.A.J., 2016. Drought in the anthropocene. *Nat. Geosci.* 9, 89–91.
- Van Loon, A.F., Kumar, R., Mishra, V., 2017. Testing the use of standardised indices and GRACE satellite data to estimate the European 2015 groundwater drought in near-real time. *Hydrol. Earth Syst. Sci.* 21, 1947–1971.
- Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A Multiscalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index – SPEI. *J. Clim.* 23, 1696–1718.
- von Asmuth, J.R., Bierkens, M.F.P., Maas, K., 2002. Transfer function-noise modelling in continuous time using predefined impulse response functions. *Water Resour. Res.* 38 (12) 23-1 23-12.
- von Asmuth, J.R., Bierkens, M.F., 2005. Modelling irregularly spaced residual series as a continuous stochastic process. *Water Resour. Res.* 41, W12404.
- Webster, R., Oliver, M.A., 1990. *Statistical Methods in Soil and Land Resource Survey*. Oxford University Press, UK.
- Webster, R., Oliver, M.A., 2007. *Geostatistics for Environmental Scientists*, second edition. John Wiley & Sons, Chichester.
- Whitehead, E., Lawrence, A.R., 2006. The Chalk aquifer system of Lincolnshire. British Geological Survey Research Report, RR/06/03.